Mechanistic studies of acetolactate decarboxylase

by

Victoria Ashley Marlow

Thesis

Submitted to the University of Warwick
for partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Molecular Organisation and Assembly in Cells

June 2013

Table of Contents

List of Figures

List of Schemes

List of Tables

Acknowledgements

Firstly, I would like to thank both my supervisors, Martin Wills and Vilmos Fülöp, for all their support and guidance throughout my postgraduate studies. They have pushed me to achieve and I am grateful to have had the opportunity to work with and learn from both of them.

It has been a pleasure to work with all members, past and present, of the Wills group and Structural Biology lab. In particular, Dr Dean Rea whose support and encouragement have been invaluable, Claire Dow for all her hard work during her mini project and Dr Katherine Abrahams for being such a fantastic friend, proof-reading my thesis so thoroughly and all the encouragement and motivation to keep me going.

I am also indebted to Alison Rodger for giving me the chance to study at MOAC, the support and motivation when it all seemed impossible and the use of her lab and advice for my CD experiments. I would also like to acknowledge the entire MOAC community, especially the 2008 cohort.

Special thanks are reserved for Sarah Shute and Anne Maynard for all their behind the scenes support. I must also thank all the support staff in the School of Life Sciences, the Department of Chemistry, the ESRF and the Diamond Light Source who have helped with various aspects of my research.

I can't forget to acknowledge the EPSRC this project would not have been possible without their funding. I am also grateful for bursaries from CCP4, the Biochemical Society and the ASBMB to attend conferences. I also want to thank Novozymes for the kind gift of purified ALDC.

It has been a challenge to finish writing my thesis since starting working for EPSRC in January and I want to thank all my colleagues within the Healthcare Technologies theme and across the organisation for all their support and for keeping me going when it seemed never

ending.

Which finally leaves me to thank my family and friends. Thank you to my parents, Brenda Marlow, Rex Marlow and Keith Coyle, my sister Lucy and my brother Josh for all your support over the years and for shaping me into be the person I am today. Thank you to Luke, Evyenia and Karen for your love and friendship. Thank you to Paul, Jennifer, Jonathan and Maggie for making me feel like part of the family and Chris for being by my side every step of the way.

Dedications

This thesis is dedicated to my late grandparents,

Ronald and Annie Paffey

Daniel and Gladys Coyle

Declarations

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. The research described in this thesis has been carried out by myself under the supervision of Prof. Vilmos Fülöp and Prof. Martin Wills. It has been composed by myself and has not been submitted in any previous application for any degree, apart from preliminary work for Sections 3.3, 5.3 and 6.3.1, which was previously submitted as part of a MSc degree. Claire Dow contributed to the work described in Section 4.3.3, which was submitted as part of Claire's MSc degree.

Parts of this thesis have been published by the author:

V. A. Marlow, D. Rea, S. Najmudin, M. Wills and V. Fülöp, 2013. "Structure and mechanism of acetolactate decarboxylase", *ACS Chemical Biology*, vol. 8, no. 10, pages 2339-2344.

Abstract

Acetolactate decarboxylase (ALDC) is a bacterial enzyme of the butanediol fermentation pathway that decarboxylates (*S*)-acetolactate into (*R*)-acetoin. Remarkably this enzyme also catalyses the decarboxylation of the opposite enantiomer, (*R*)-acetolactate, to give the same product, (*R*)-acetoin. It is unusual for an enzyme to convert racemic substrate into an enantiomerically pure product. This unusual stereochemical control has led to extensive study of the ALDC mechanism and the hypothesis that ALDC catalyses the rearrangement of (*R*)-acetolactate into (*S*)-acetolactate prior to decarboxylation.

The research presented in this thesis sought to gain insight into the molecular mechanism of the ALDC catalysed reaction by using a combination of kinetic and structural techniques. *Bacillus subtilis alsD* encoding ALDC was cloned into an expression vector and a series of active site mutants were prepared. The activity of mutant AlsD were determined using a circular dichroism based assay, which identified that the two active site glutamates and a basic residue are required for catalysis.

A series of chiral transition state analogues were prepared in a two-step synthesis to give enantiomerically enriched 2,3-dihydroxy-2-methylbutanoic acid in reasonable yields. Three of the compounds were identified as competitive inhibitors, co-crystallised with *Bacillus brevis* ALDC and structures solved to 1.1-1.6 Å. These structures, coupled with inhibition studies and site-directed mutagenesis, provide an improved understanding of the molecular processes involved in the stereoselective decarboxylation of acetolactate. A mechanism for the transformation of each enantiomer of acetolactate is proposed.

Abbreviations

^{13}C	Carbon-13
AADase	Acetoacetate decarboxylase
ACMSD	α-Amino-β-carboxymuconic-ε-semialdehyde decarboxylase
ALDC	Acetolactate decarboxylase
ALS	Acetolactate synthase
AR	Acetoin reductase
AU	Asymmetric unit
B. brevis	*Bacillus brevis*
BCAA	Branched chain amino acid
Bis-Tris	Bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methane
bp	Base pair
B. subtilis	*Bacillus subtilis*
CCD	Charged coupled device
CD	Circular dichroism
cryo-EM	Cryo-Electron Microscopy
$\Delta\varepsilon$	Molar Circular Dichroism

$(DHQD)_2PHAL$	Hydroquinidine-1,4-phthalazinediyl ether
$(DHQ)_2PHAL$	Hydroquinine-1,4-phthalazinediyl ether
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
E	Enzyme
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
ee	Enantiomeric excess
EG	Ethylene glycol
EI	Enzyme-Inhibitor complex
EIS	Enzyme-Inhibitor-Substrate complex
ES	Enzyme-Substrate complex
ESI	Electrospray Ionisation
ESRF	European Synchrotron Radiation Facility
GC	Gas chromatography
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRMS	High resolution mass spectrometry
I	Inhibitor
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl B-D-1-thiogalactopyranoside
IR	Infrared
$k_{\mathbf{cat}}$	Turnover number

K_i	Inhibition constant
K_m	Michaelis-Menten constant
L. lactis	*Lactobacillus lactis*
LB	Luria Bertani
MES	2-(N-morpholino)ethanesulfonic acid
min	minute
MME	Monomethyl ether
MS	Mass spectrometry
MW	Molecular weight
NAD^+	Nicotinamide adenine dinucleotide (oxidised)
NADH	Nicotinamide adenine dinucleotide (reduced)
NMR	Nuclear Magnetic Resonance
OD_{600}	Optical density at 600 nm
OMP	Orotidine 5'-phosphate
OMPDC	Orotidine 5'-phosphate decarboxylase
P	Product
PAGE	Polyacrylamide gel elctrophoresis
PBS	Phosphate buffered saline
PCET	Proton coupled electron transfer
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PLP	Pyridoxal phosphate

psi	Pound per square inch
PVDF	Polyvinylidene difluoride
RMSD	Root mean square deviation
rpm	Revolutions per minute
rt	Room temperature
S	Substrate
SAD	Single wavelength anomalous dispersion
S. aureus	*Staphylococcus aureus*
SCOP	Structural class of protein
SDS	Sodium Dodecyl Sulphate
sec	Second
SOC	Super Optimal broth with Catabolite repression
TAE	Tris Acetate Ethylenediaminetetraacetic acid
TEMED	Tetramethylethylenediamine
TLC	Thin layer chromatography
TMAO	Trimethylamine N-oxide
TPP	Thiamin pyrophosphate
Tris	Tris (hydroxymethyl) aminomethane
UMP	Uridine 5'-monophosphate
UV	Ultraviolet
$V_{\mathbf{max}}$	Maximal enzyme rate
VP	Voges-Proskauer

Chapter 1

Introduction

1.1 Enzymes as biological catalysts

Enzymes are a subset of proteins that catalyse the biochemical reactions necessary for life; they are present in every cell of all living organisms. They are able to accelerate the rate of a reaction by lowering the activation energy required for that reaction to take place. Enzymes achieve this by either stabilising the high energy transition state of the reaction or destabilising the ground state. The rate of an enzyme catalysed reaction can be several orders of magnitude greater than the uncatalysed reaction. This high efficiency is achieved through a strong affinity for the incredibly short-lived transition state structure [1–6]. Although enzymes are large macromolecules, catalysis usually occurs in a small region of the protein known as the active site. The specificity of the enzyme-catalysed reaction is determined by the arrangement of amino acids within the active site. Some enzymes show absolute specificity and only turn over a single substrate, other enzymes exhibit broader substrate specificities whereby a family of related substrates can be turned over. The chiral nature of an enzyme active site confers high stereospecificity of the reaction catalysed; an enzyme can

distinguish between enantiomers or confer chirality on a prochiral substrate. There are many commercial uses for enzymes outside of their physiological roles, for example enzymes are used in food-processing [7], brewing [8], biofuel production [9], detergents [10] and synthetic chemistry as biocatalysts [11].

Detailed understanding of an enzyme mechanism requires information that cannot be gained from a single technique. Structural techniques such as X-ray crystallography and nuclear magnetic resonance spectroscopy (NMR) can give information about the arrangement of amino acid residues in the enzyme active site, molecular biology techniques can be used to produce active site mutants and spectroscopic techniques can be used to measure enzyme kinetics. Techniques such as these, used in parallel often provides sufficient information to describe the mechanism of an enzyme.

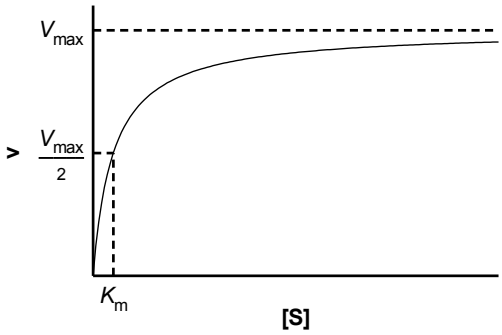
1.2 Enzyme kinetics

Studying the kinetics of an enzyme-catalysed reaction by monitoring the reaction rate can provide insight into the enzyme mechanism, provide information about how small molecules inhibit the enzyme and how mutation of active site residues affects activity.

Typical studies of enzyme kinetics involve measuring the initial reaction rate over a range of substrate concentrations, which can be plotted to give a hyperbolic curve, an example plot is shown in Figure 1.1. Figure 1.1 shows that the rate of an enzyme-catalysed reaction increases with increasing substrate concentration and then levels off as it reaches a maximal rate, V_{max}. Also shown in Figure 1.1 is the Michaelis-Menten constant, K_{m}, which is the substrate concentration required for the enzyme to operate at half the maximal rate.

The observed behaviour can be explained by the Michelis-Menten model of enzyme kinetics;

Figure 1.1: The Michaelis-Menten model of enzyme kinetics



A schematic of an enzyme catalysed reaction. When the substrate concentration [S] is plotted against the initial rate (v), the rate increases until the enzyme is saturated and a maximal rate is achieved (V_{max}). The michaelis-menten constant (K_m) is the substrate concentration required for the enzyme to work at half the maximal rate (V_{max}). The relationship between v and [S] is described by the Michaelis-Menten equation (Equation 1.13).

the model was developed by Leonor Michaelis and Maud Menten in 1913 [12,13]. The following sections will describe the derivation of the Michaelis-Menten equation, the importance of the kinetic parameters and the study of enzyme inhibition.

1.2.1 Derivation of the Michaelis-Menten equation

The model shown in Scheme 1.1 states that enzyme (E) combines with substrate (S) to form an enzyme-substrate complex (ES). The enzyme-substrate complex can then dissociate back to (E) and (S) or proceed to form product (P).

Scheme 1.1: Reaction scheme for an enzyme catalysed reaction

$$[\mathrm{E}] + [\mathrm{S}] \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [\mathrm{ES}] \xrightarrow{k_2} [\mathrm{E}] + [\mathrm{P}]$$

This scheme shows the simplest model of enzyme kinetics: free enzyme [E] and [S] can combine to form an enzyme-substrate complex [ES] with a rate constant of k_1. The [ES] complex can dissociate to back to [E] and [S] with a rate constant of k_{-1} or react to form [E] and product [P] with a rate constant of k_2.

According to Scheme 1.1 the rate of P formation is proportional to the concentration of the ES complex as described by Equation 1.1

$$\frac{d[\text{P}]}{dt} = v = k_2[\text{ES}] \tag{1.1}$$

The concentration of ES complex is dependent on the rate of formation of ES (Equation 1.2) and the rate of dissociation of ES (Equation 1.3).

$$\frac{d[\text{ES}]}{dt} = k_1[\text{E}][\text{S}] \tag{1.2}$$

$$\frac{d[\text{ES}]}{dt} = k_{-1}[\text{ES}] + k_2[\text{ES}] \tag{1.3}$$

Under steady state conditions, when [S]»[E], the concentration of ES complex is assumed to remain constant therefore the rate of formation of ES equals the rate of dissociation as shown in Equation 1.4.

$$k_1[\text{E}][\text{S}] = k_{-1}[\text{ES}] + k_2[\text{ES}] \tag{1.4}$$

Rearranging Equation 1.4 gives Equation 1.5.

$$\frac{[\text{E}][\text{S}]}{[\text{ES}]} = \frac{k_{-1} + k_2}{k_1} \tag{1.5}$$

The Michaelis constant K_m is defined by Equation 1.6

$$\frac{k_{-1} + k_2}{k_1} = K_m \tag{1.6}$$

Substituting K_m into Equation 1.5 gives Equation 1.7

$$\frac{[\text{E}][\text{S}]}{[\text{ES}]} = K_m \tag{1.7}$$

The free enzyme concentration equals the total enzyme concentration (E_T) minus the concentration of enzyme-substrate complex, as shown in Equation 1.8.

$$[E] = [E_T] - [ES] \tag{1.8}$$

Substituting Equation 1.8 into Equation 1.7 gives Equation 1.9.

$$\frac{([E_T] - [ES])[S]}{[ES]} = K_m \tag{1.9}$$

Rearranging Equation 1.9 in terms of [ES] gives Equation 1.10.

$$[ES] = \frac{[E_T][S]}{K_m + [S]} \tag{1.10}$$

Substituting Equation 1.10 into Equation 1.1 gives Equation 1.11.

$$v = \frac{k_2[E_T][S]}{K_m + [S]} \tag{1.11}$$

k_2 is also known as the kinetic constant k_{cat}, which can be defined by Equation 1.12

$$k_{cat} = \frac{V_{max}}{E_T} \tag{1.12}$$

Substituting Equation 1.12 into Equation 1.11 gives the Michaelis-Menten equation (Equation 1.13).

$$v = \frac{V_{max}[S]}{K_m + [S]} \tag{1.13}$$

1.2.2 Determination of kinetic constants

There are three important kinetic parameters that can be determined from a typical enzyme activity assay; V_{max}, K_m and k_{cat}. Enzyme-catalysed reactions get saturated at high substrate concentrations and reach a maximal rate known as V_{max}. At sufficiently high [S], K_m becomes negligible and the Michaelis-Menten equation (Equation 1.13) cancels to give Equation 1.14.

$$v = \frac{V_{max}[S]}{[S]} = V_{max} \tag{1.14}$$

The turnover number, k_{cat}, can be determined from V_{max} if the total enzyme concentration is known (Equation 1.12). k_{cat} is defined as the maximum number of enzyme catalysed reactions per active site per unit time and is usually given units of s^{-1}. The greater the value of k_{cat} the faster the enzyme. The value of k_{cat} is limited to around 10^8-10^9 s^{-1} as when an enzyme operates at near perfect efficiency it is limited by the rate of diffusion of substrate [14].

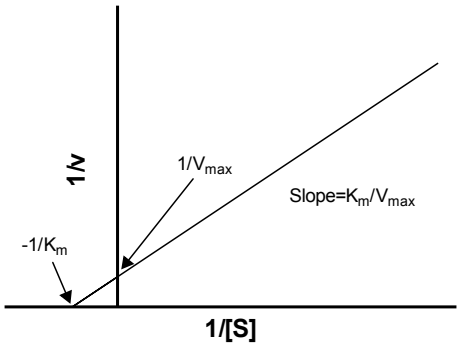
The Michaelis-Menten constant K_m has units of concentration and is the concentration of substrate required to give half the maximal rate. This can be demonstrated by substituting K_m=[S] into the Michaelis-Menten equation.

$$v = \frac{V_{max}K_m}{K_m + K_m} = \frac{V_{max}K_m}{2K_m} = \frac{V_{max}}{2} \tag{1.15}$$

In equation 1.6, K_m is defined as $(k_{-1} + k_2)/k_1$. Under conditions where $k_1 >> k_2$ then K_m is equal to the dissociation constant of the ES complex and is a measure of enzyme substrate binding. A high K_m indicates weak binding whereas a low K_m indicates strong binding.

The term k_{cat}/K_m is known as the specificity constant and is a measure of catalytic efficiency.

Figure 1.2: Lineweaver-Burk plot of enzyme kinetics



The Lineweaver-Burk plot is a double reciprocal plot of 1/[S] versus 1/v, which linearises the Michaelis-Menten equation. The kinetic parameters can be determined from the plot, the y intercept is $1/V_{max}$, the x intercept is $-1/K_m$ and the slope is K_m/V_{max}. This type of plot is useful as a graphical representation but is prone to errors if used to calculate kinetic parameters.

Enzyme efficiency is limited by two factors: the binding interaction between enzyme and substrate and the turnover rate of an enzyme. The specificity constant takes both factors into account. The larger the value of the specificity constant the more efficient the enzyme.

The kinetic constants can be determined by plotting substrate concentration versus rate and then fitting the data to the Michaelis-Menten equation using non-linear regression. Prior to the computer age, kinetic constants had to determined graphically. Due to the difficulty in determining values from a hyperbolic curve, methods were developed to represent the data linearly. The most commonly used graphical representation, the Lineweaver-Burk plot [15], is a double reciprocal of $1/v$ versus 1/[S]. Figure 1.2 shows a schematic of the Lineweaver-Burke plot, the y intercept is $1/V_{\mathrm{max}}$, the x intercept is $-1/K_{\mathrm{m}}$ and the slope is $K_{\mathrm{m}}/V_{\mathrm{max}}$. The Lineweaver-Burk plot was very useful for rapidly determining kinetic parameters in the laboratory; however it has limitations. This method is unreliable as it introduces error, as all the data points at higher substrate concentrations are compressed into a small region putting more weight on the points at lower concentration. In addition, small errors in rate measurements are magnified when reciprocals are taken, leading to errors into the determination of

kinetic constants.

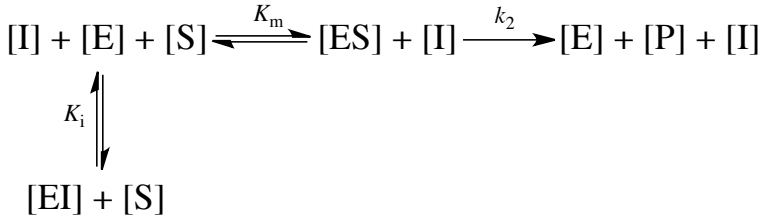
1.2.3 Enzyme inhibition

There are four types of reversible inhibition: competitive, non-competitive, mixed and un-competitive inhibition. It is possible to determine the mode of inhibition and the inhibitor dissociation constant (K_{i}) by measuring substrate concentration versus initial rate curves in the presence of several concentrations of inhibitor [16].

1.2.3.1 Competitive inhibition

A competitive inhibitor reversibly binds to the same site as the substrate. The inhibition can be overcome by using very high concentrations of substrate. In the presence of a competitive inhibitor the apparent K_{m}, $K_{\mathrm{m}}^{\mathrm{app}}$, increases but V_{max} remains constant.

Scheme 1.2: Reaction scheme for competitive inhibition



The reaction scheme of an enzyme catalysed reaction in the presence of a competitive inhibitor. As in an uninhibited reaction, [E] and [S] can combine to form [ES] with an equilibrium constant of K_m. The [ES] complex can react to form [E] and product [P] with a rate constant of k_2 but in addition [E] can combine with inhibitor [I] to form an unreactive [EI] complex with a dissociation constant of K_i.

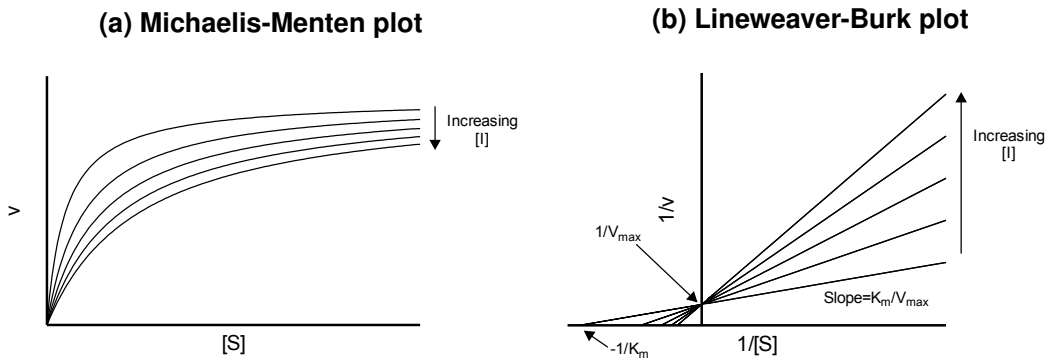
A competitive inhibitor follows the model shown in Scheme 1.2. The free enzyme can form either an enzyme-inhibitor complex (EI) or an ES complex. In the presence of a competitive inhibitor, the Michaelis-Menten equation is modified as shown in Equation 1.16 and $K_{\mathrm{m}}^{\mathrm{app}}$ is defined in Equation 1.17.

$$v = \frac{V_{\text{max}}[\text{S}]}{K_{\text{m}}^{\text{app}} + [\text{S}]} \tag{1.16}$$

$$K_{\text{m}}^{\text{app}} = K_{\text{m}}(1 + \frac{[\text{I}]}{K_{\text{i}}}) \tag{1.17}$$

K_{i} is the dissociation constant for the EI complex and a lower value indicates a stronger inhibitor. Figure 1.3a shows a typical Michaelis-Menten plot for a competitive inhibitor, the V_{max} is unchanged but $K_{\text{m}}^{\text{app}}$ increases. Figure 1.3b shows the same data on a Lineweaver-Burk plot, as [I] increases the slope increases but the intercept on the y-axis is unchanged indicating an increase in $K_{\text{m}}^{\text{app}}$ but a constant V_{max}.

Figure 1.3: Competitive inhibition

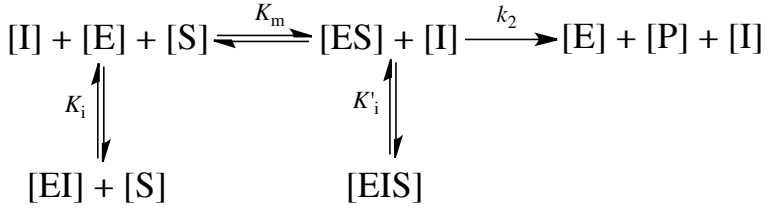


Competitive inhibition is characterised by an increase in the apparent K_m and no change in the V_{max} as shown in a typical (a) v versus [S] plot. (b) This can also been shown on a Lineweaver-Burk plot, where competitive inhibition is characterised by an increasing slope and a fixed y intercept.

1.2.3.2 Non-competitive Inhibition

Non-competitive inhibition occurs when an inhibitor binds to both E and ES with equal affinity as shown in Scheme 1.3.

Scheme 1.3: Reaction scheme for non-competitive inhibition



The reaction scheme of an enzyme catalysed reaction in the presence of a non-competitive inhibitor. As in a competitively inhibited reaction, [E] and [S] can combine to form [ES] with an equilibrium constant of K_m. The [ES] complex can react to form [E] and product [P] with a rate constant of k_2 and [E] can combine with inhibitor [I] to form an unreactive [EI] complex with a dissociation constant of K_i. In addition the [ES] complex can also combine with [I] to form an unreactive [EIS] complex with a dissociation constant of K_i.

When K_i=K'_i, the Michaelis-Menten equation can be modified as shown in Equation 1.18 and $V_\mathrm{max}^\mathrm{app}$ is defined by Equation 1.19.

$$v = \frac{V_\mathrm{max}^\mathrm{app}[\mathrm{S}]}{K_\mathrm{m} + [\mathrm{S}]} \tag{1.18}$$

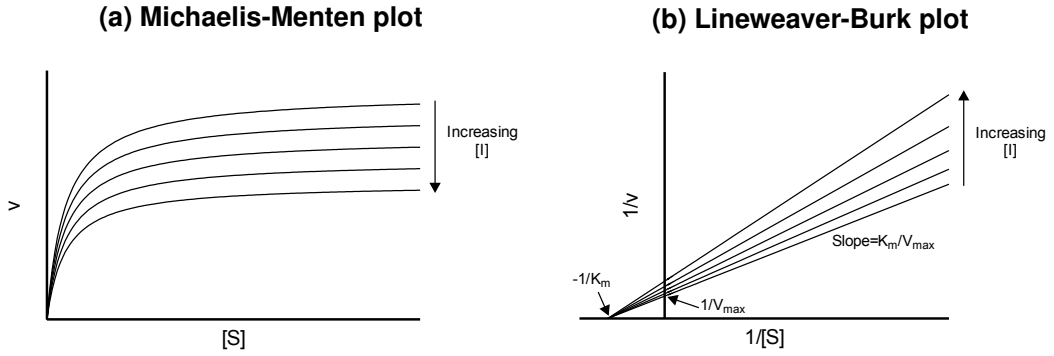
$$V_\mathrm{max}^\mathrm{app} = \frac{V_\mathrm{max}}{(1 + \frac{[\mathrm{I}]}{K_\mathrm{i}})} \tag{1.19}$$

Figure 1.4a shows [S] versus v curves for non-competitive inhibition, it causes a decrease in the apparent V_max, $V_\mathrm{max}^\mathrm{app}$, and no change in the K_m. The same data on a Lineweaver-Burk plot is shown in Figure 1.4b. The lines all intercept on the x-axis indicating a constant K_m but the increasing slope and increasing y-axis intercepts indicate a decrease in $V_\mathrm{max}^\mathrm{app}$.

1.2.3.3 Mixed Inhibition

A mixed inhibitor is a type of non-competitive inhibitor where the affinity of I is different for the free enzyme and the ES complex, $K_\mathrm{i} \neq K'_\mathrm{i}$. The inhibition can be thought of as a mixture of competitive and uncompetitive and it is characterised by a decrease in $V_\mathrm{max}^\mathrm{app}$ and can

Figure 1.4: Non-competitive inhibition



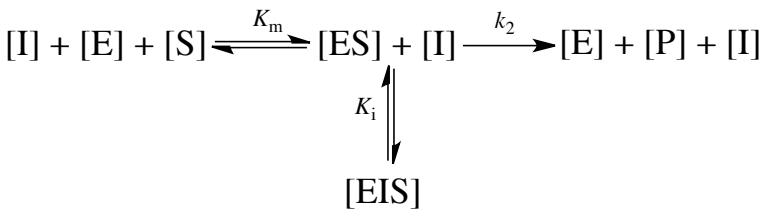
Non-competitive inhibition is characterised by a decrease in the apparent V_{max} and no change in the K_m as shown in a typical (a) v versus [S] plot. (b) This can also been shown on a Lineweaver-Burk plot, where non-competitive inhibition is characterised by an increasing slope, an increasing y intercept and a fixed x intercept.

increase or decrease the K_{m}. The Lineweaver-Burk plot for mixed inhibition is similar to that of non-competitive inhibition however the lines do not intersect on the x-axis, instead they intersect either above or below it.

1.2.3.4 Uncompetitive inhibition

Scheme 1.4 shows the model for uncompetitive inhibition, the inhibitor does not bind to the free enzyme and only binds to the ES complex.

Scheme 1.4: Reaction scheme for uncompetitive inhibition



The reaction scheme of an enzyme catalysed reaction in the presence of an uncompetitive inhibitor. Unlike the other types of inhibitor, an uncompetitive inhibitor cannot bind to [E], it can only bind to the [ES] complex to form an unreactive [EIS] complex with a dissociation constant of K_i.

The effect of this type of inhibition is to decrease both the apparent V_{max} and the apparent K_{m}. The Michaelis-Menten equation for uncompetitive inhibition is shown in Equation 1.20 and $V_{\text{max}}^{\text{app}}$ and $K_{\text{m}}^{\text{app}}$ are defined in Equations 1.21 and 1.22 respectively.

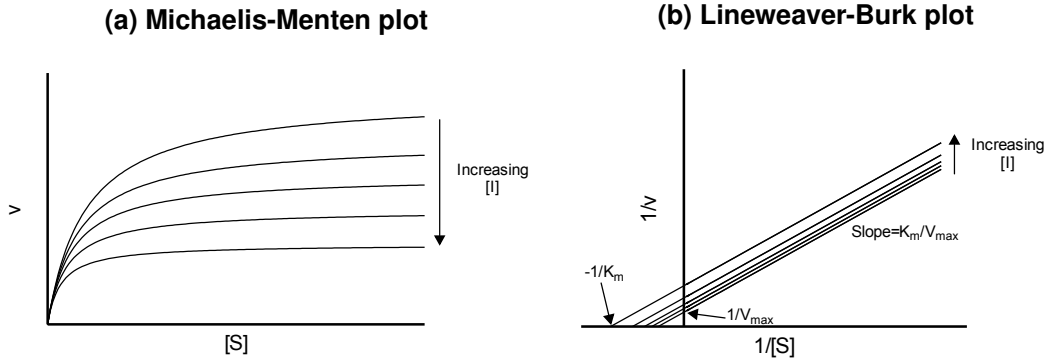
$$v_0 = \frac{V_{\text{max}}^{\text{app}}[\text{S}]}{K_{\text{m}}^{\text{app}} + [\text{S}]} \tag{1.20}$$

$$V_{\text{max}}^{\text{app}} = \frac{V_{\text{max}}}{(1 + \frac{[\text{I}]}{K_{\text{i}}})} \tag{1.21}$$

$$K_{\text{m}}^{\text{app}} = \frac{K_{\text{m}}}{(1 + \frac{[\text{I}]}{K_{\text{i}}})} \tag{1.22}$$

Figure 1.5a shows example [S] versus v curves for an uncompetitive inhibitor and Figure 1.5b shows the same data on a Lineweaver-Burk plot. The parallel lines on the Lineweaver-Burk plot indicate that increasing inhibitor concentrations cause a decrease in both $V_{\text{max}}^{\text{app}}$ and $K_{\text{m}}^{\text{app}}$.

Figure 1.5: Uncompetitive inhibition



Uncompetitive inhibition is characterised by a decrease in the apparent V_{max} and a decrease in the apparent K_m as shown in a typical (a) v versus [S] plot. (b) This can also been shown on a Lineweaver-Burk plot, where uncompetitive inhibition is characterised by a series of parallel lines with no change in the slope but an increase in the y intercept.

1.3 X-ray Crystallography

There are three main techniques used for imaging biological macromolecules at near atomic resolution: X-ray crystallography, NMR and cryo-electron microscopy (cryo-EM). As of the 5th March 2013 there were 88714 biological macromolecular structures deposited in the RCSB protein data bank (www.rcsb.org), of these 78162 structures were solved by X-ray crystallography, 9869 of the structures were solved by NMR, 516 structures were solved by cryo-EM and the remaining 167 were solved by other techniques such as neutron diffraction, fibre diffraction, electron crystallography, powder diffraction and solution scattering.

The three major techniques all have their advantages and disadvantages. Cryo-EM is commonly used to look at large macromolecular complexes, the samples are cryo-cooled rapidly to prevent the formation of vitreous ice so the samples are in a frozen yet hydrated state. NMR spectroscopy is the technique of choice for looking at dynamic and flexible proteins as structures can be solved in solution, however the size is usually limited to small or medium proteins. X-ray crystallography is the technique of choice for determining atomic resolution structures as it can provide high resolution structures, however it is necessary to grow protein crystals, which is challenging and is not possible for all proteins. X-ray crystal structures give a snapshot of the protein in a conformation that was constructive to crystal growth and it is important to remember that proteins are dynamic and only a single conformation is observed for an X-ray crystal structure.

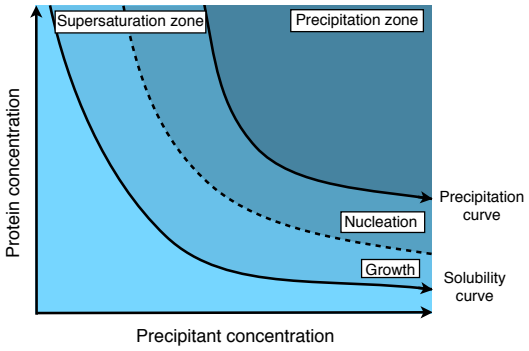
It is not possible to use a light microscope to visualise protein structures, as the resolving power of visible light is limited to its wavelength which is around 0.5 μm, meaning mitrochondria are the smallest cellular structures that can be visualised. To be able to resolve at atomic resolution, it is necessary to use radiation with a wavelength of 1 Å (0.1 nm). X-rays fall within this region of the electromagnetic spectrum, however it is not possible to build an

X-ray microscope to image at atomic resolution. In a light microscope, the lens focuses the light to create an image of the object but it is not possible to create an X-ray lens as the refractive index of X-rays is close to 1 when passing through all materials. Instead we can collect the diffracted X-rays; traditionally diffracted X-rays were collected on film but more recently they are collected using either imaging plates, charge-coupled devices (CCDs) or solid state detectors. The structure of the target molecule can be solved indirectly from the diffraction pattern through the use of computers that can implement Fourier Transform techniques.

1.3.1 Crystallisation

The use of protein crystals in X-ray diffraction studies is necessary to amplify the signal, as a single protein molecule would result in diffraction that is too weak to measure. A protein crystal is a regular array of repeating units that act as a diffraction grating to amplify the diffraction signal. The process of growing crystals is challenging, crystallisation occurs when the concentration of protein in the solution is greater than its limit of solubility. Meeting this criterion alone will most likely lead to the protein precipitating out of solution. If the conditions are carefully controlled the protein can enter a supersaturated phase in which crystal nucleation can occur; once crystal nuclei have formed crystals can grow. Figure 1.6 shows the important stages of crystal growth on a phase diagram of protein concentration versus precipitant concentration. To grow single diffraction quality crystals, it is important that the conditions between nucleation and growth are met. The optimal conditions for nucleation and crystal growth are not the same as there is typically a higher energy barrier required to nucleate a crystal which requires greater levels of supersaturation [17,18]. The two most common methods of growing protein crystals are vapour diffusion experiments, known as the hanging drop and sitting drop methods [19]. In both methods, a droplet containing protein

Figure 1.6: Phase diagram of crystallisation



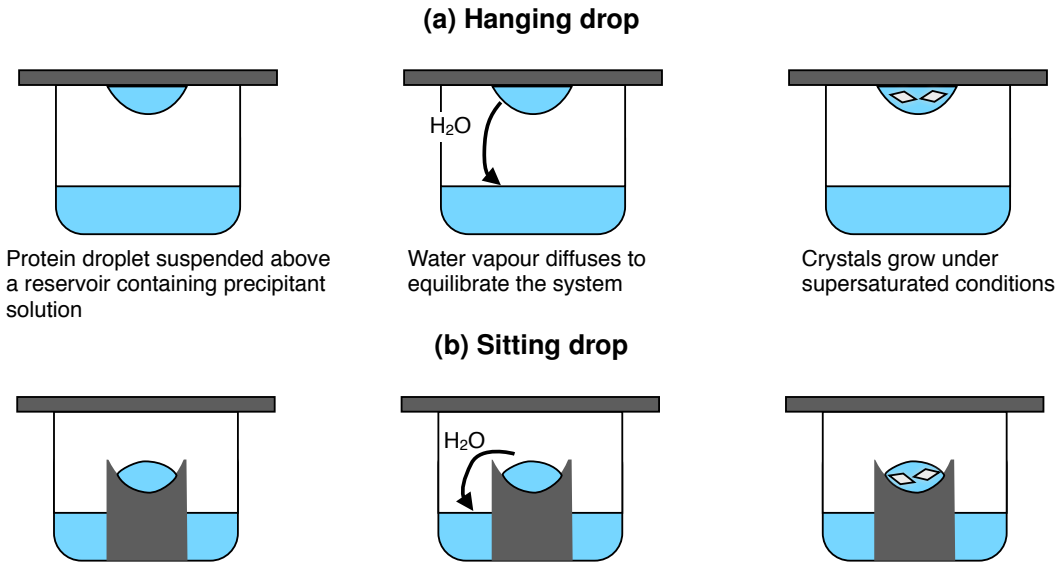
The phase diagram shows that at low protein and precipitant concentrations the protein is soluble, whereas at high concentrations of protein and precipitant, the protein precipitates out of solution. Under narrow conditions a supersaturation region occurs where protein is no longer soluble but will not precipitate. It is in this region that protein crystallisation occurs. Protein crystallisation occurs in two stages, nucleation and growth. Protein crystallisation methods create a trajectory through the phase diagram that hopefully promotes crystallisation.

and precipitant is allowed to equilibrate with a larger reservoir containing precipitant solution. The concentration of precipitant is higher in the reservoir than the droplet so water vapour diffuses from the droplet into the reservoir. As water vapour diffuses out of the droplet, both the protein and precipitant increase in concentration. This allows crystal nucleation to occur and as crystals begin to grow the protein concentration in the droplet lowers. In the sitting drop method the principle of crystallisation is the same, the only difference is the placement of the droplet, instead of being suspended on a cover slip above the reservoir it is sitting in a well suspended above the reservoir as shown in Figure 1.7.

Achieving protein crystallisation can be described as a "black art" as it is a case of trial and error. Screening of protein crystallisation conditions used to be a laborious task but with advances in automated crystallisation robots setting up a 96 well commercial screen can take a matter of minutes and use very low volumes of sample.

Once a "hit" has been found it often requires optimisation to yield the best quality crystals.

Figure 1.7: Vapour diffusion method of protein crystallisation



The vapour diffusion method of protein crystallisation can be arranged in two different experimental set ups: (a) hanging drop vapour diffusion and (b) sitting drop vapour diffusion. In both methods a droplet containing protein is suspended above a reservoir containing precipitant solution in a sealed system. The precipitant concentration is higher in the reservoir solution than in the protein droplet, which causes water vapour to diffuse from the droplet to the reservoir to equilibrate the system. As water vapour diffuses from the protein droplet the concentration of protein and precipitant in the droplet rises causing the solution to reach supersaturation where crystal nucleation and growth can occur.

Optimisation can involve varying the pH, precipitant and/or protein concentration and adding compounds known to enhance crystallisation [19].

1.3.2 Data collection

Once suitable crystallisation conditions have been found it is necessary to subject the crystals to diffraction experiments. Data collection is typically carried out at 100 K, as at these low temperatures the effect of radiation damage on crystals is slowed, meaning more data can be collected before the diffraction of a crystal is affected. Carrying out data collection at 100 K requires crystals to be cryo-protected to prevent the formation of ice crystals during cryo-cooling. Typically small molecules like glycerol or ethylene glycol are used as cryoprotectants although there are many other cryoprotectants available.

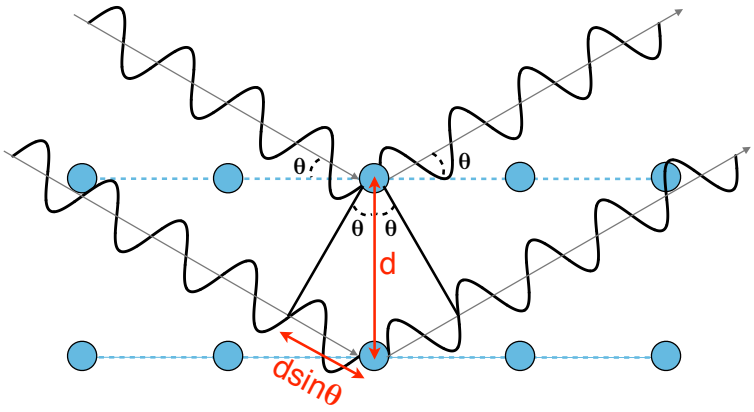
Crystals are formed of repeating unit cells arranged in a lattice and as such can only exist in a limited number of crystal symmetries. There are a total of 230 different crystal symmetries based on rotation, reflection and inversion symmetry. Protein crystals can only adopt 65 of the possible symmetries as the chiral nature of proteins make reflection and inversion symmetries impossible. The data that is collected from an X-ray diffraction experiment is a series of diffraction images, each one taken as the crystal is rotated slightly. Information about the crystal symmetry is present within the diffraction pattern; the symmetry of intensities with the diffraction pattern is known as Laue symmetry. For the 65 space groups only 11 Laue symmetries are observed, due to the presence of mirror symmetry in the diffraction image. Although mirror symmetry cannot be observed in the crystals themselves, Friedel's Law states that the intensity of reflection (hkl) equals that of reflection (-h-k-l). The space group cannot be determined solely from the diffraction image, however looking for symmetry in intensities and systematic absences of reflections can provide information about the

rotation and translation symmetry of the crystal. A typical data collection experiment begins with taking two diffraction images at 90° apart and these images are then indexed using software such as iMOSFLM [20], which identifies a list of space groups based on the observed diffraction pattern and predicts a data collection strategy. Using a data collection strategy allows the best quality data to be obtained whilst the crystal is in good condition [21].

Not all X-rays that pass through a crystal produce a reflection in the diffraction image, it is only when the conditions of Bragg's law are met that a reflection is observed. Bragg's law is defined by Equation 1.23 and its derivation is described by Figure 1.8. For a reflection to be observed in the diffraction image, the scattered X-rays must be in phase.

$$n\lambda = 2d\sin\theta \tag{1.23}$$

Figure 1.8: Derivation of Bragg's law



The blue lines represent planes of atoms that are distance d apart. Bragg's law states that reflected X-rays must remain in phase to interfere constructively and hence result in observable reflections . The wavelength of the X-rays is λ and the angle of incidence θ is equal to the angle of reflection. The second wave travels a greater distance than the first wave. This additional distance can be determined by applying trigonometry. Bragg's law states that this distance 2dsinθ must be equal to a whole number of wavelengths (nλ) for the waves to stay in phase. If the conditions of Bragg's law are not met the waves are out of phase and cancel out due to destructive interference.

1.3.3 Data processing

After a data collection experiment, a series of diffraction images need to be processed into an electron density map. The first step in this process is to index the images to generate a list of reflection positions and intensities. The next step is to produce an electron density map. Electron density can be calculated using Fourier Transform if the amplitudes and phases of the scattered waves are known. The Fourier Transform is given by Equation 1.24a.

$$\rho(x,y,z) = \frac{1}{V}\sum_{hkl} F_{hkl}[\cos(-2\pi(hx+ky+lz)) + i\sin(-2\pi(hx+ky+lz))] \tag{1.24a}$$

$$F_{hkl} = |F_{hkl}|(\cos\phi_{hkl} + i\sin\phi_{hkl}) \tag{1.24b}$$

F_{hkl} is the structure factor which includes the terms for amplitude $|F_{hkl}|$ and phase ϕ as shown in Equation 1.24b. The data generated from a diffraction experiment contains reflection intensities which are proportional to amplitude (Equation 1.25), however the phase information is lost.

$$I_{hkl} \propto |F_{hkl}|^2 \tag{1.25}$$

The inability to measure the phase information directly from an X-ray diffraction experiment is known as the phase problem.

Methods to determine the phases can be classified into two categories: experimental phasing and molecular replacement [22]. Molecular replacement is the technique of choice as it is much simpler than experimental phasing which requires further data collection. If a structure is known that has homology to the unsolved structure, then the phases from this known structure can be used as an initial solution that gets improved in an iterative process. If there is no structure with significant homology for a molecular replacement solution to be found

then experimental phasing methods must be used. Experimental phasing falls broadly into two categories: isomorphous replacement and anomalous scattering. Isomorphous replacement uses heavy atoms to provide phase information. Heavy atoms scatter X-rays strongly and the presence of a heavy atom causes changes in intensity of some reflections in the diffraction pattern. If the location of the heavy atom is known then the phase problem can be solved. Anomalous scattering can also be used to solve the phase problem. If the X-ray wavelength used in the data collection is tuned to the absorption edge of an anomalously scattering atom then the intensity changes in the diffraction pattern can be used to provide phase information.

Once the phases are known, the Fourier Transform can be solved and an electron density map calculated. If the quality of the initial map is insufficient to start model building then the quality can be improved using density modification methods such as solvent flattening, histogram matching and density averaging of non-crystallographic symmetry related molecules. Once the map quality is improved then model building can begin. After an initial model is built, it goes through an iterative process of refinement and manual building with the aim to improve the fit of the model to the data. The quality of the model is assessed by the R-factor, which is defined in Equation 1.26.

$$R = \frac{\sum ||F_{obs}| - |F_{calc}||}{\sum |F_{obs}|} \tag{1.26}$$

As the model is improved through refinement, the R value decreases. Lower R values indicate better structures and typically an R-factor of around or less than 20% has been deemed sufficient for publication. Another criteria that is used to judge the quality of the model is R_{free} [23]. After data processing a portion of the data are set aside and not used in the refinement. This test set is used to prevent over-fitting of the model to the data. If there is an

improvement in the R-factor but not in R_{free} then over-fitting has occurred.

1.4 Decarboxylases

Enzymes can be classified into six classes based on the reactions that they catalyse: oxidoreductases, transferases, hydrolases, lyases, isomerases and ligases. Acetolactate decarboxylase (EC 4.1.1.5) falls in the fourth catergory, lyases and in particular, a carboxylyase (or decarboxylase), which cleaves a carbon-carbon bond with the removal of carbon dioxide. In general for decarboxylation to occur there needs to be a β-electron sink; some substrates naturally incorporate such an electron sink but others utilise active site residues or cofactors. Decarboxylases can be broadly grouped into three categories based on the nature of the cofactor used: cofactor independent, organic cofactor dependent or metal ion dependent.

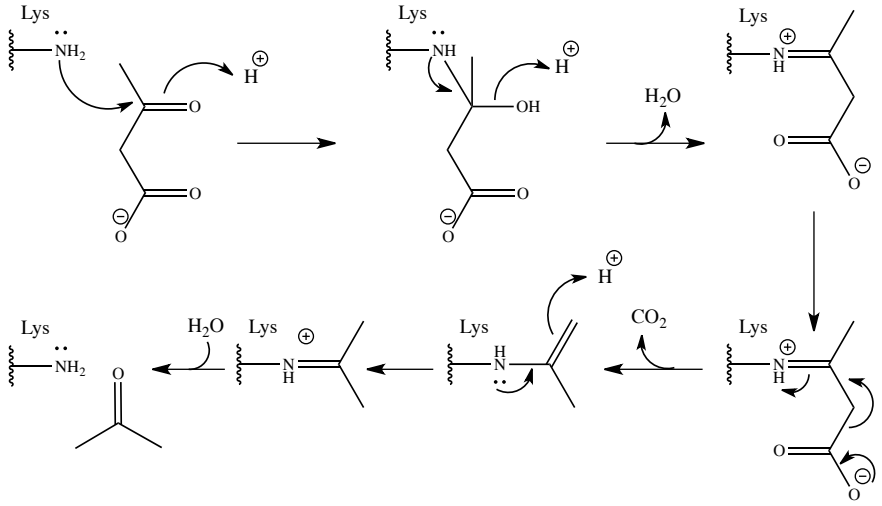
1.4.1 Cofactor independent decarboxylation

Acetoacetate decarboxylase (AADase) and orotidine 5'monophosphate decarboxylase (OMPDC) are examples of decarboxylases that do not require a cofactor [24]. Acetoacetate is a β-keto acid and the ketone group acts as a natural β-electron sink. Although AADase does not require a cofactor it does use a Schiff base mechanism of catalysis. AADase is an example of an enzyme that undergoes dramatic pKa perturbations. The pKa of lysine is usually around 10.6. However the pKa of the ε-aminium group in Lys115 is around 6, which is unusually low. Initially it was proposed that the low pKa was due to a positive electrostatic interaction between Lys115 and neighbouring residue Lys116, resulting in the deprotonation of Lys115 [25]. However, eludication of the crystal structure of AADase has shown that the pKa perturbation occurs due to a desolvation effect of placing Lys115 within a hydrophobic

pocket [26]. The NH_2 group of Lys115 catalyses a nucleophilic attack on acetoacetate forming a Schiff base. The resulting β-iminium ion undergoes decarboxylation as shown in Figure 1.9. The residue numbering refers to AADase from *Clostridium acetobutylicum*.

OMPDC catalyses the decarboxylation of orotidine 5'-monophosphate (OMP) to uridine 5'-monophosphate (UMP), which is the final step in the pyrimidine biosynthetic pathway. It is one of the most proficient enzymes known with a rate enhancement of 17 orders of magnitude without the requirement of any cofactors [27]. The mechanism of OMPDC is rather unusual as it does not require a β-electron sink. The current mechanism is still disputed in the literature, however it is proposed to occur by a combination of substrate destabilisation due to the electrostatic repulsion from the close proximity of an active site aspartate residue with the substrate carboxylate and transition state stabilisation of the carbanion by a conserved lysine residue [28,29].

Figure 1.9: Mechanism of acetoacetate decarboxylase



The active site lysine residue acts as a nucleophile, attacking the substrate, acetoacetate, to form a Schiff base. This promotes decarboxylation as the negative charge is stabilised by the formation of an enamine intermediate. The product, acetone, is released by tautomerisation of the enamine intermediate followed by hydrolysis.

1.4.2 Organic cofactor dependent decarboxylation

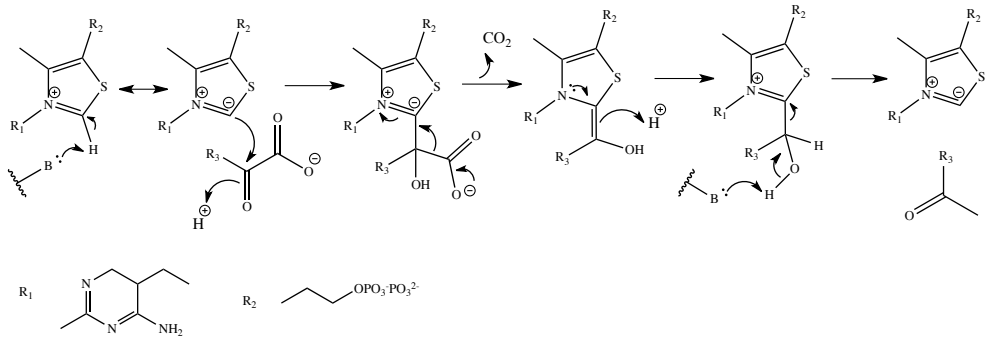
A wide variety of organic cofactors are utilised by decarboxylases but they generally activate the substrate by providing a β-electron sink to the substrate. The three most commonly required organic cofactors are thiamin pyrophosphate (TPP), pyridoxal phosphate (PLP) and a pyruvoyl group [24]. The general mechanisms for TPP, PLP and pyruvoyl dependent decarboxylation is shown in Figure 1.10 and some examples are described below.

TPP is derived from vitamin B1 and is used as a cofactor by many enzymes. It consists of a pyrimidine ring connected to a thiazolium ring which is connected to a pyrophosphate group. The most common substrates for TPP dependent decarboxylases are α-ketoacids. Yeast pyruvate decarboxylase is an example of a TPP dependent decarboxlase [30]. Pyruvate does not have a natural β-electron sink but it reacts with the TPP ylide as shown in Figure 1.10a to form a pyruvate-TPP adduct. The pyruvate-TPP adduct is activated for decarboxylation as the thiazolium ring of TPP acts as a β-electron sink. Decarboxylation occurs leaving an enamine intermediate, which gets protonated to release acetaldehyde and regenerate the TPP ylide [24].

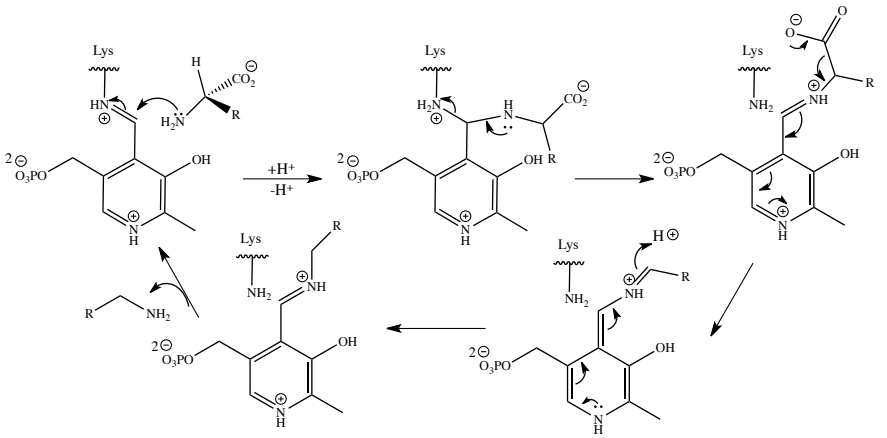
PLP is an organic cofactor that is derived from vitamin B6. It is a cofactor that is found covalently linked to enzymes that require it. It is used in a variety of reactions involving aminated compounds. PLP dependent decarboxylases typically catalyse decarboxylation of amino acids and the general mechanism is shown in Figure 1.10b. PLP dependent decarboxylases such as ornithine decarboxylase [31] or aromatic L-amino acid decarboxylase [32] undergo a Schiff base mechanism. In the resting state, PLP is found covalently bound to an active site lysine residue through an imine bond. The first step in catalysis is a transimination to form a Schiff base intermediate aldimine between PLP and substrate amino group. The formation of the substrate-PLP aldimine intermediate activates the substrate for decarboxylation by

Figure 1.10: Mechanisms of cofactor dependent decarboxylases

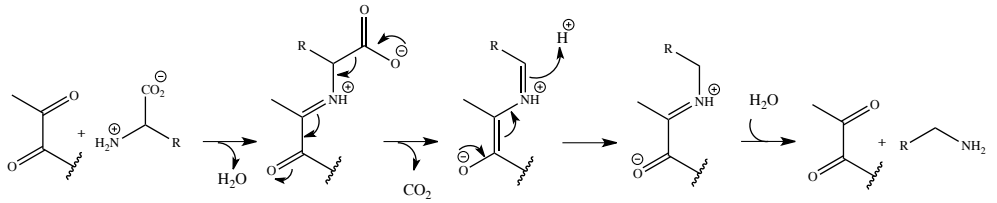
(a) TPP dependent decarboxylation



(b) PLP dependent decarboxylation



(c) pyruvoyl dependent decarboxylation



General mechanisms of cofactor dependent mechanisms. (a) Thiamin pyrophosphate (TPP) is often used in the decarboxylation of α-keto acids; the thiazolium ring of the cofactor acts as a β-electron sink to activate the substrate for decarboxylation. (b) Pyridoxal phosphate (PLP) is a common cofactor used by amino acid decarboxylases. PLP forms a Schiff base intermediate to promote decarboxylation. (c) Alternatively some amino acid decarboxylases use an active site pyruvoyl group as a cofactor. These enzymes also form a Schiff base intermediate to promote catalysis.

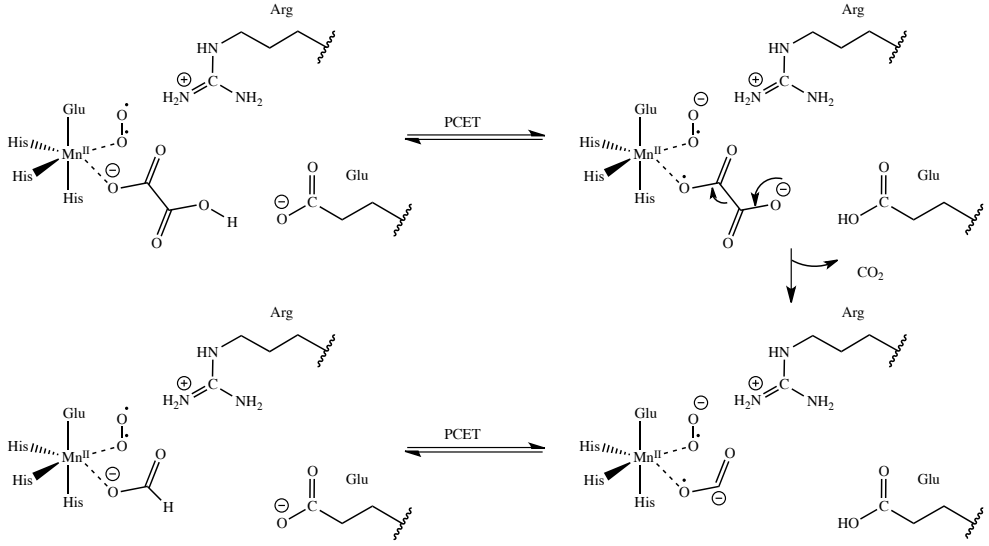
providing an electron sink. The decarboxylated intermediate is protonated at C_α to form an imine intermediate. Transimination of PLP back to Lys releases the amine product [24].

Some amino acid decarboxylases use an alternative cofactor: an active site pyruvoyl group that is covalently bound via an amide linkage in the enzyme active site. Examples include arginine decarboxylase [33], *S*-adenosylmethionine decarboxylase [34] and bacterial histidine decarboxylase [35]. As shown in Figure 1.10c, the first step in catalysis is the formation of a Schiff base between the amino group of the substrate and the active site pyruvoyl group. The acyl-carbonyl group of the pyruvoyl-substrate imine intermediate acts as the electron sink and the subsequent steps of the decarboxylation mechanism proceed in a homologous manner as for the PLP dependent decarboxylases [24].

1.4.3 Metal ion dependent decarboxylation

Many decarboxylases require metal ions as catalytic cofactors and can be classified as oxidative or non-oxidative depending on their requirement for oxygen. The proposed mechanism of Mn^{2+} dependent, oxalate decarboxylase is shown in Figure 1.11. In the resting state, the Mn^{2+} in the active site is coordinated to three histidines and either an aspartate or glutamate residue and two water molecules. The substrate oxalate displaces one of the water molecules to coordinate monodentately to the metal and the second water molecule is displaced by dioxygen. In the first step, oxalate is activated for decarboxylation by a proton coupled electron transfer (PCET) to form an oxalate radical. The activated oxalate radical decarboxylates, forming a formate radical intermediate. A second PCET generates formate. Saylor et al. [36] propose that oxygen rather than manganese acts as the electron sink due to the failure to observe catalytically competent Mn^{3+} or Mn^{4+} species. Other metal dependent oxidative decarboxylases include Mn^{2+} dependent 4-oxalocronate decarboxylase and

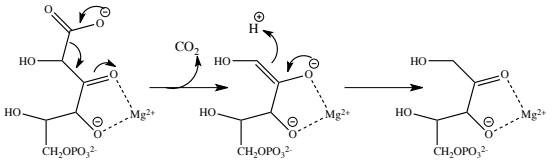
Figure 1.11: Mechanism of oxalate decarboxylase



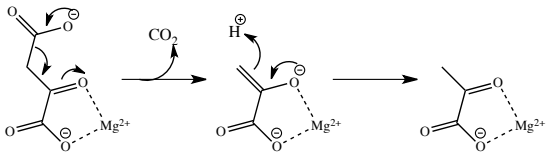
Oxalate decarboxylase is a Mn^{2+} and O_2 dependent decarboxylase. The mechanism is proposed to occur through a proton coupled electron transfer to generate an oxalate radical intermediate. The intermediate undergoes decarboxylation and a further proton coupled electron transfer to release the product and regenerate the catalyst [36].

Figure 1.12: Mechanisms of Mg^{2+} dependent decarboxylases

(a) Dehydro-L-gluconate 6-phosphate decarboxylase



(b) Oxaloacetate decarboxylase



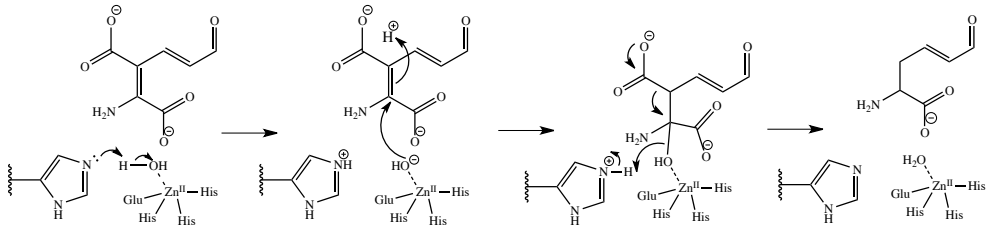
(a) Dehydro-L-gluconate 6-phosphate decarboxylase is a non-oxidative Mg^{2+} dependent decarboxylase. Substrate coordinates to the active site metal through two O-Mg coordination bonds, an enolate intermediate stabilises the negative charge of the intermediate. (b) Oxaloacetate decarboxylase is also a Mg^{2+} dependent decarboxylase which acts in an analogous way to dehydro-L-gluconate 6-phosphate.

Fe^{2+} dependent decarboxylases such as gallic acid decarboxylase and the α-ketoglutarate-dependent dioxygenases [24].

Dehydro-L-gluconate 6-phosphate decarboxylase [37] and oxaloacetate decarboxylase [38] are examples of non-oxidative metal dependent decarboxylases. The proposed mechanisms for both enzymes are shown in Figure 1.12a and 1.12b, respectively. Both substrates co-ordinate to Mg^{2+} (a non-transition metal cofactor) bidentately. The decarboxylation of both compounds is stabilised by formation an enolate intermediate which is then protonated and product released. The metal ion acts to stabilise the intermediates in both cases [24].

α-Amino-β-carboxymuconic-ε-semialdehyde decarboxylase (ACMSD) is a transition metal dependent non-oxidative decarboxylase requiring an active site Zn^{2+} for catalysis. The active site of ACMSD contains Zn^{2+} coordinated to three histidines, an aspartate residue and a water molecule. The proposed mechanism for ACMSD is shown in Figure 1.13, where a fourth active site histidine is proposed to deprotonate the active site water. The metal bound hydroxide then nucleophillically attacks the substrate at the β-position relative to the leaving carboxylate. The OH then acts as an electron sink for decarboxylation to occur [39,40].

Figure 1.13: Mechanism of ACMSD



α-Amino-β-carboxymuconic–semialdehyde decarboxylase is a Zn^{2+} dependent decarboxylase. The metal acts to coordinate a water molecule which is deprotonated by an active site histidine to form hydroxide. The hydroxide ion acts as a nucleophile to attack the substrate. This intermediate is decarboxylated and the hydroxide ion acts to stabilise the negative charge.

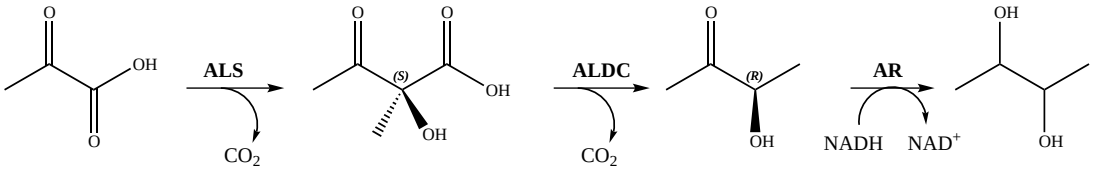
1.5 Acetolactate decarboxylase

1.5.1 Biological role of ALDC

Acetolactate decarboxylase (ALDC) is an enzyme involved in the butanediol biosynthetic pathway, which converts pyruvate **1** into 2,3-butanediol **2** in three steps as shown in Figure 1.14. In the first step, catalysed by acetolactate synthase, two molecules of pyruvate **1** are combined to form acetolactate **3**. Then ALDC catalyses the decarboxylation of acetolactate **3** to acetoin **4** with the release of CO_2. Finally, acetoin reductase converts acetoin **4** into 2,3-butanediol **2** [41]. The pathway is specific to bacteria and over 40 bacterial species encoding an ALDC have been identified [42]. The metabolic function of the pathway is unknown, however it has been hypothesised to be a method of switching metabolism from acid production to the production of neutral compounds. Studies of *Klebsiella terrigena* and *Enterobacter aerogenes* have shown that the three genes are encoded in a single operon that is activated in oxygen limiting conditions and at approximately pH 6 [43,44].

ALDC produced by *Lactococcus lactis* is unlike other butanediol producing bacteria as it has

Figure 1.14: Butane-2,3-diol fermentation pathway



*The butane-2,3-diol fermentation pathway is a bacterial pathway that consists of three enzymes: acetolactate synthase (ALS), acetolactate decarboxylase ALDC and acetoin reductase (AR). Two molecules of pyruvate are combined by ALS to form (S)-acetolactate with the release of CO_2. Acetolactate is decarboxylated by ALDC to form (R)-**4** with the release of another molecule of CO_2. Finally AR reduces acetoin to butane-2,3-diol in an NADH dependent manner.*

been shown to have a dual role, not only is it involved in the biosynthesis of butanediol, it is also involved in the regulation of the branched chain amino acids (BCAA). Acetolactate is a precursor in the synthesis of the BCAA, leucine and valine but these are typically produced by a different pathway operating under different conditions. In the case of *L. lactis* the gene encoding ALDC is located within the BCAA operon and it has been demonstrated *in vitro* that ALDC is under allosteric control of leucine and is believed to have a role in regulation of branched amino acid synthesis [45].

The ALDC from *Bacillus brevis* is rather unusual in that it appears to be an exoenzyme as the gene encodes a signal peptide for external transport. The role of *B. brevis* ALDC as an extracellular enzyme is inconsistent with the physiological roles so far identified and is unknown [46].

1.5.2 Industrial applications of ALDC

1.5.2.1 Brewing

Acetolactate decarboxylase has been used in the brewing industry for many years to shorten the maturation step. Brewing consists of two major steps: fermentation and maturation. During fermentation, alcohol is produced by the anaerobic metabolism of sugars by yeast. The maturation step occurs after the main fermentation, as it is required to develop the flavour of the beer. Traditionally in brewing the maturation step could take up to 12 weeks [46–48]. This step is necessary because during beer production diacetyl is formed by the spontaneous oxidative decarboxylation of acetolactate. Diacetyl conveys an "off" taste to beer in concentrations above 0.1 ppm [49]. Acetolactate is produced by the yeast as a byproduct of branched chain amino acid biosynthesis. The maturation step is required to allow any acetolactate to oxidatively decarboxylate to diacetyl which can then be converted to acetoin

by diacetyl reductase. Without the maturation step the diacetyl concentration in the beer rises giving it an undesirable taste. Godtfredsen et al. [48] first suggested the use of ALDC to increase the efficiency of the brewing process by shortening the rate-limiting beer maturation step. Many methods of applying ALDC to the brewing process have been developed; it can be applied as an exogenous free enzyme [42,48–50], an encapsulated enzyme which is recoverable [51] or incorporated into the yeast genome for continuous expression during the fermentation process [52]. The use of ALDC in this way has been shown to significantly increase the production rate of beer without any adverse flavours on the final product [48].

1.5.2.2 Synthesis of enantiomerically pure diols

ALDC has been used to generate a whole-cell biocatalyst for the production of enantiomerically pure 2,3-butanediol from glucose. Yan et al. [53] introduced three genes encoding acetolactate synthase, acetolactate decarboxylase and acetoin reductase into *E. coli* on a plasmid under the control of pL*lac*O1 promoter. As ALDC produces exclusively *R*-acetoin they were able to generate biocatalysts for the production of either *(2R,3R)* or *(meso)*-butane-2,3-diol with enantiomeric purity over 99% by using acetoin reductases with different specificity.

1.5.3 Mechanistic studies of ALDC

ALDC is a unique enzyme that exhibits an unusual level of stereochemical control: it can decarboxylate both enantiomers of acetolactate **3** yet only gives a single enantiomer of the decarboxylation product, (*R*)-acetoin (*R*)-**4**. The unusual stereochemistry of this reaction has led to the extensive study of the ALDC mechanism [54–60].

Initially, it was thought that (*S*)-acetolactate (*S*)-**3** was the sole substrate for ALDC. The decarboxylation of (*R*)-acetolactate (*R*)-**3**, observed by Hill et al. [55] and Crout et al. [56], was

originally attributed to a slow racemisation of (*R*)-acetolactate (*R*)-**3** into (*S*)-acetolactate (*S*)-**3** by the known base catalysed mechanism [54,61]. Observations that acetolactate **3** racemises in dilute alkali [61] were shown to occur through a tertiary ketol rearrangement with carboxylate migration [54]. The mechanism of rearrangement was determined by a sophisticated NMR study of ^{13}C labelled acetolactate. Doubly labelled [1,3-^{13}C]-acetolactate was prepared and treated with dilute sodium hydroxide solution in the presence of an excess of unlabelled acetolactate **3** and monitored by NMR. Once the reaction had reached equilibrium, NMR signals identified a mixture of [1,3-^{13}C]-acetolactate [1,3-^{13}C]-**3** and [1,2-^{13}C]-acetolactate [1,2-^{13}C]-**3**, indicating the rearrangement had occurred via an intramolecular carboxylate migration [54].

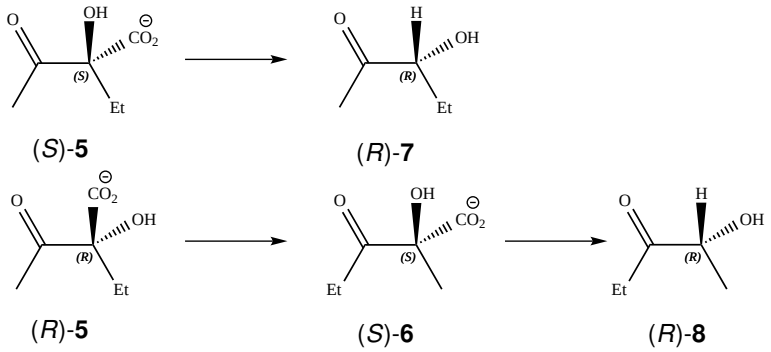
The pH dependence of this rearrangement was determined by monitoring acetohydroxybutyrate **5** at four different pH values. When acetohydroxybutyrate **5** undergoes carboxylate migration, it is converted to the isomeric product 2-hydroxy-2-methyl-3-oxopentanoate **6**; these two isomeric ketols (shown in Figure 1.15) can be distinguished by NMR. At pH 14, the signals for both isomers were equal, indicating complete racemisation within 3 hours. At pH 12.9, the ratio of acetohydroxybutyrate **5** to 2-hydroxy-2-methyl-3-oxopentanoate **6** was 2:1 after 7 hours. At pH 12.5 and 12.1 no racemisation had occurred within 7 hours, indicating that carboxylate migration only occurs under alkaline conditions with a pH > 12.5 [56].

Figure 1.15: Base-catalysed rearrangement of acetohydroxybutyrate

5 **6**

*Acetohydroxybutyrate **5** undergoes a base-catalysed rearrangement by carboxylate migration at pH > 12.5 to give the isomeric compound 2-hydroxy-2-methyl-3-oxopentanoate **6**.*

Figure 1.16: ALDC-catalysed decarboxylation of acetohydroxybutyrate

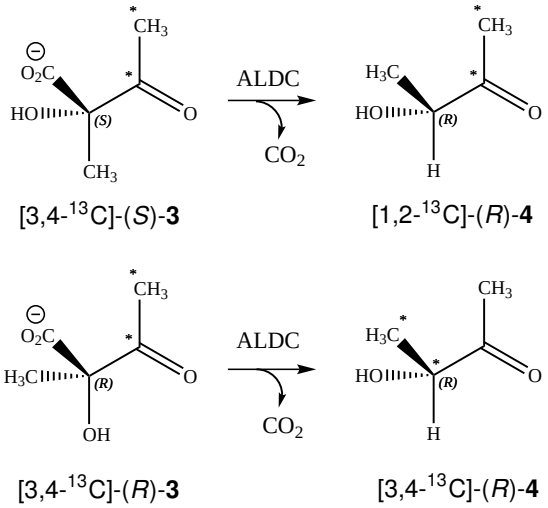


*ALDC catalyses the decarboxylation of (S)-acetohydroxybutyrate (S)-**5** to give (R)-3-hydroxypentan-2-one (R)-**7**. Whereas (R)-acetohydroxybutyrate (R)-**5** is hypothesised to undergo an enzyme-catalysed rearrangement to give (S)-2-hydroxy-2-methyl-3-oxopentanoate (S)-**6** which is then decarboxylated to give (R)-2-hydroxypentan-3-one (R)-**8**.*

This observation led to the hypothesis that the slow decarboxylation of (*R*)-acetolactate (*R*)-**3** that had been attributed to base catalysed rearrangement, may in fact have been enzyme catalysed [62].

Crout et al. [58] tested this hypothesis by studying the ALDC catalysed decarboxylation of acetohydroxybutyrate **5**. They proposed that if ALDC was capable of catalysing the carboxylate migration rearrangement then (*S*)-acetohydroxybutyrate (*S*)-**5** would be decarboxylated, preferentially, to give (*R*)-3-hydroxypentan-2-one (*R*)-**7** and that (*R*)-acetohydroxybutyrate (*R*)-**5** would be converted to the structural isomer (*S*)-2-hydroxy-2-methyl-3-oxopentanoate (*S*)-**6** by carboxylate migration, which could be decarboxylated to give (*R*)-2-hydroxypentan-3-one (*R*)-**8**, as shown in Figure 1.16. The results agreed with their hypothesis, the decarboxylation of racemic acetohydroxybutyrate (±)-**5** was monitored by NMR, initially a rapid decarboxylation was observed with signals for substrate disappearing as signals for the expected product **7** appeared. After approximately 50% of the substrate had disappeared, a second slower decarboxylation occurred as signals corresponding to the formation of **8** appeared. Although the stereochemistry of the compounds was not identified, products **7** and

Figure 1.17: ^{13}C labelling studies of the ALDC catalysed reaction



[3,4-^{13}C]-(*S*)-**3** [1,2-^{13}C]-(*R*)-**4**

[3,4-^{13}C]-(*R*)-**3** [3,4-^{13}C]-(*R*)-**4**

The ALDC catalysed reaction was studied using ^{13}C-labelled acetolactate and monitored by NMR, positions of the ^{13}C are indicated by an asterix (=^{13}C). ALDC converts [3,4-^{13}C]-(S)-acetolactate [3,4-^{13}C]-(S)-**3** into [1,2-^{13}C]-(R)-acetoin [1,2-^{13}C]-(R)-**4** but converts [3,4-^{13}C]-(R)-acetolactate [3,4-^{13}C]-(R)-**3** into [3,4-^{13}C]-(R)-acetoin [3,4-^{13}C]-(R)-**4**.*

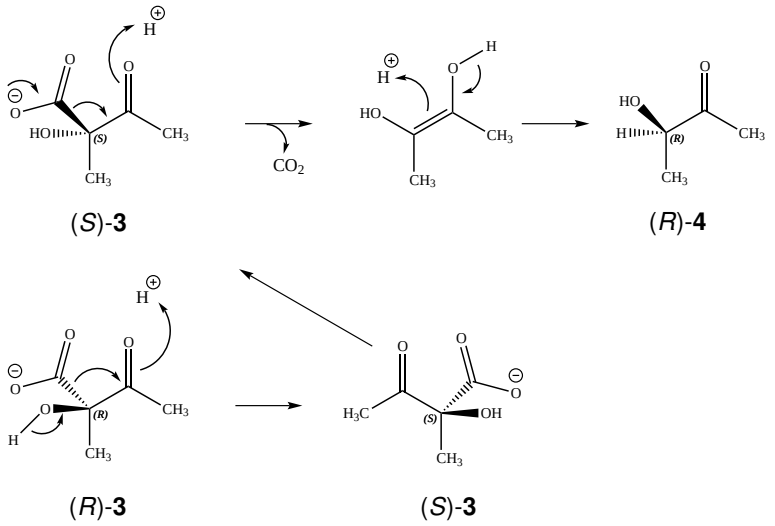
8 were shown to have enantiomeric excesses (ee) of 93 and 95%, respectively. The authors also confirmed that ALDC was able to decarboxylate racemic acetolactate to give exclusively (*R*)-acetoin with an ee of >98%.

A further study [59] showed the same rearrangement for acetolactate. The decarboxylation of doubly labelled [3,4-^{13}C]-acetolactate [3,4-^{13}C]-**3** by ALDC was monitored by NMR. Initially, signals appeared that corresponded to [1,2-^{13}C]-acetoin [1,2-^{13}C]-**4**, however once the reaction reached approximately 40% completion, signals corresponding to [3,4-^{13}C]-acetoin [3,4-^{13}C]-**4** began to appear. This is shown diagrammatically in Figure 1.17. This experiment confirmed that the preferential substrate for ALDC is (*S*)-acetolactate (*S*)-**3**, which is decarboxylated with inversion of configuration and (*R*)-acetolactate (*R*)-**3** is decarboxylated more slowly with protonation occurring on the carbon originally containing the ketol group in the substrate. The protonation step has been shown to occur stereospecifically by studying D/H exchange of acetoin [57]. When acetolactate decarboxylase was incubated with racemic

acetoin in D_2O, only 50% of the acetoin underwent deuterium exchange of the methine proton as shown by ^{1}H-NMR. Vibrational CD confirmed that the deuterated component was (*R*)-acetoin and both these results indicated that the protonation step controls the stereochemistry of the ALDC product.

In conclusion, it can be summarised, that prior to the work described in this thesis, the mechanism of ALDC was hypothesised to occur in two distinct stages, as shown in Figure 1.18. The preferred substrate of ALDC, (*S*)-acetolactate, is decarboxylated via an enediol intermediate that is stereospecifically protonated with inversion of configuration. The alternative substrate, (*R*)-acetolactate, is hypothesised to undergo an enzyme-catalysed carboxylation migration rearrangement to give (*S*)-acetolactate, which can be decarboxylated. Although the mechanism has been extensively studied, little is known about the enzymatic control of the reaction.

Figure 1.18: Proposed mechanism of ALDC



ALDC is proposed to decarboxylate (S)-acetolactate with inversion of configuration to generate (R)-acetoin. However the enzyme is also proposed to decarboxylate (R)-acetolactate, by catalysing a tertiary ketol rearrangement with carboxylate migration forming (S)-acetolactate which can then be decarboxylated.

1.5.4 Structure of ALDC

ALDC is not only an interesting enzyme because of its unusual stereochemical control, it is structurally very interesting as it has a novel protein fold. There are two X-ray crystallography structures of ALDC, from *Bacillus brevis* and *Staphylcoccus aureus*.

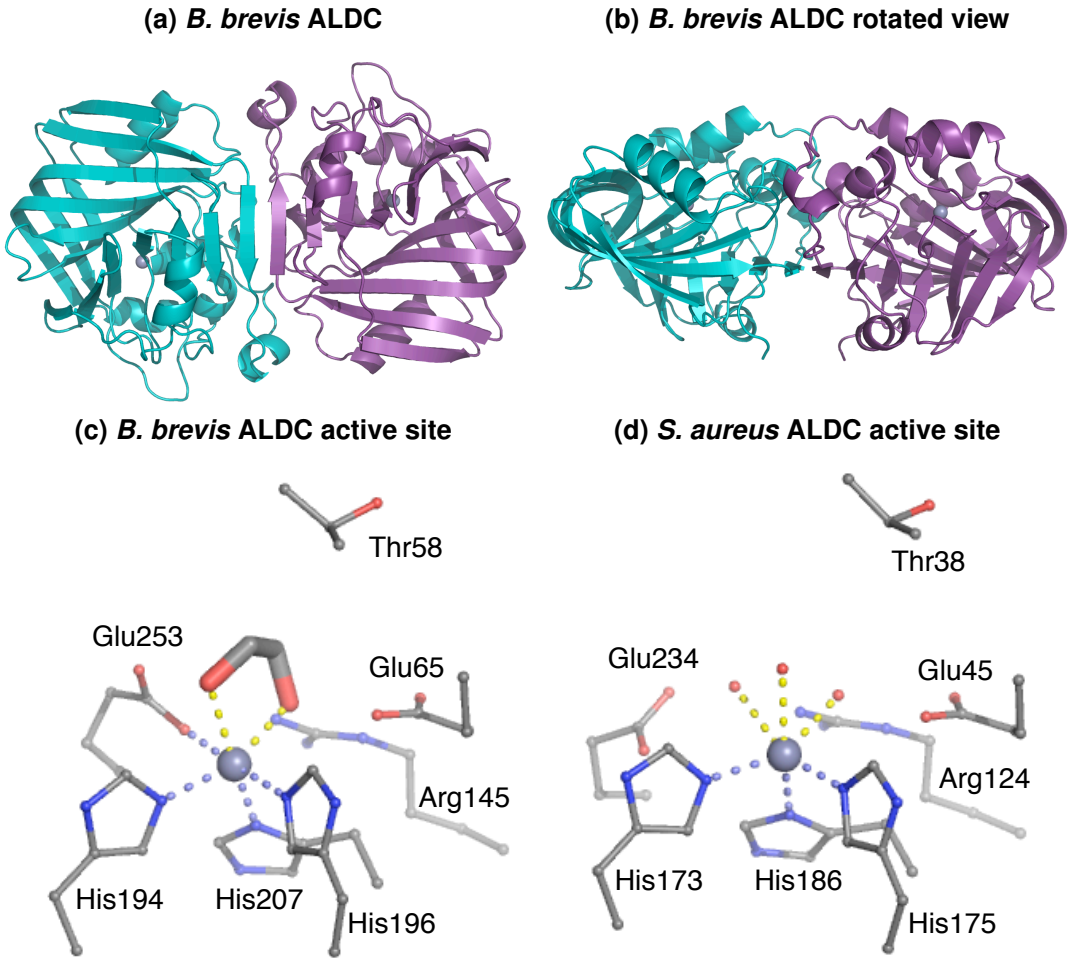
The structure of *B. brevis* ALDC has been previously solved by the Fülöp group at the University of Warwick. Some preliminary crystallisation statistics were published [63] and the structure was subsequently solved by anomalous scattering but the structure (4BT2) was only published recently in conjunction with the results presented in this thesis [64]. *B. brevis* ALDC was over-expressed in *Bacillus subtilis* strain JA222, which contains multiple copies of the *aldB* gene. ALDC was purified from the fermentation supernatant by cation exchange chromatography followed by size-exclusion chromatography and crystallised by the hanging drop vapour diffusion method. Three different crystal forms were described: type 1 belonged to space group $P4_222$, contained 1 molecule per asymmetric unit (AU) and diffracted to 2.2 Å; type 2 crystals belonged to space group $P3_121$ or $P3_221$ with 1 molecule in the AU and diffracted to 2 Å; type 3 crystals belonged to space group I222 with 3 molecules in the AU and diffracted to 2.7 Å [63]. This work was continued and the type 2 crystals were further optimised and the structure solved by anomalous scattering, the crystal belonged to space group $P3_221$ with unit cell parameters a=46.95, b=46.95, c=198.90 Å, diffracting to 1.1 Å resolution and the refinement statistics were R_{value}=0.189 and R_{free}=0.208 [64].

Since the *B. brevis* structure was solved, the structure of ALDC from *S. aureus* has been solved and deposited in the protein data bank (1XV2). This structure has been solved by the Midwest Centre for Structural Genomics; it was described as a hypothetical protein with similarity to acetolactate decarboxylase. Although it was deposited in 2004 there has been no follow up publication. *S. aureus* ALDC was recombinantly expressed in *E. coli* BL21

(DE3) and crystallised by the hanging drop vapour diffusion method. Crystals belonged to space group $P2_1$ with unit cell dimensions of a=48.43, b=160.01, c=68.06 Å and diffracted to 2 Å. Four molecules were present in the AU and the refinement statistics were R_{value}=0.191 and R_{free}=0.236.

Both structures are highly similar with a root mean square deviation (RMSD) of 0.7 Å between the two structures, the two proteins have a 26% sequence identity and a 38% sequence similarity. The ALDC fold is a novel protein fold; it belongs to its own SCOP (structural classification of proteins) family, the alpha-acetolactate decarboxylase-like family. Figure 1.19a and 1.19b show the structures of the *B. brevis* ALDC dimer from two different views. The structure of *S. aureus* ALDC is included in Appendix A. ALDC is a mixed α/β protein that forms a homo-dimer with two-fold symmetry. Each monomer contains two domains, the N-terminal domain contains a 7 strand mixed β-sheet, which extends to form a 14 strand β-sheet upon dimerisation. The C-terminal domain contains a β-barrel structure that is formed from a six strand antiparallel β-sheet that folds over to form the barrel structure. The active site is located at the interface between the two domains. In the *B. brevis* structure, shown in Figure 1.19c, a Zn^{2+} ion is coordinated to three highly conserved histidines (194,196 and 207) and a glutamate (253) from the C-terminal tail. The structure also shows ethylene glycol (EG) bound to the metal ion. EG was present in the crystallisation conditions and is a product analogue. Also in close proximity are Thr58, Arg145 and Glu65, which are also highly conserved and possibly play a role in catalysis or stabilisation of intermediates. The active site of the *S. aureus* ALDC, shown in Figure 1.19d is highly similar containing Zn^{2+} coordinated to three histidines (173,175 and 186), also in close proximity are Glu45, Thr38, Arg124. However, in this structure Glu234 from the C-terminal tail is not coordinated to the metal, indicating that there may be some flexibility in the C-terminus which may be important in catalysis.

Figure 1.19: Structure of ALDC



(a) The structure of B. brevis ALDC at 1.1 Å resolution. Monomers are shown in cyan and purple, the two N-terminal domains form a β sheet at the dimer interface. (b) The dimer viewed from a 90° rotation shows the β barrel from the C-terminal domain. (c) The B. brevis active site contains Zn^{2+} (grey sphere) coordinated to three histidines, a glutamate from the C-terminal tail and product analogue ethylene glycol. (d) The S. aureus active site contains Zn^{2+} (grey sphere) coordinated to three histidines and three water molecules. The C-terminal glutamate is shifted away from the metal. The structure of the S. aureus dimer is shown in Appendix A.

These two structures have identified ALDC as a metalloprotein, and the location of the active site is known, however the molecular mechanism of catalysis was still unknown at the outset of this project. The amino acids in the active site that are likely to be involved in catalysis have been identified but it is unclear how the two enantiomers of the substrate bind in the active site. The mode of binding of ethylene glycol indicates that acetolactate may bind via two oxygen coordination bonds.

1.6 Thesis motivation

ALDC is a remarkable enzyme that is able to convert a racemic substrate into an enantiomerically pure product. There is potential for this enzyme to be developed as a biocatalyst for the synthesis of enantiomerically enriched diols. However, before this long term aim can be achieved it is necessary to understand the mechanism of ALDC at the molecular level. Extensive studies by David Crout have established the scheme of the ALDC catalysed reaction but little is known about how ALDC confers the high stereochemical control at the molecular level. The crystal structure of ALDC identified it as a Zn^{2+} dependent decarboxylase and located the active site and several highly conserved amino acids but it is still unclear what residues are important for catalysis.

The overall aim of this project was to gain insight into the molecular mechanism of acetolactate decarboxylase by using a combination of kinetic and structural techniques. This will be achieved by generating ALDC structures complexed to chiral transition state analogues and by developing an assay for ALDC activity to study the inhibition kinetics and activity of active site mutants. The combination of these approaches will give insights into how different substrates interact with the active site, identify which residues participate in catalysis and propose a mechanism for ALDC.

Chapter 2

Methods and experimental

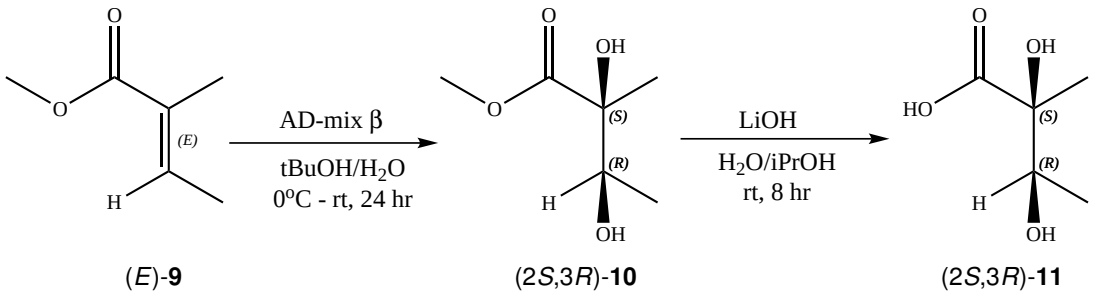
2.1 Synthesis of substrates and inhibitors of acetolactate decarboxylase

Reagents and solvents used in this study were supplied by Sigma Aldrich, TCI, VWR or Fischer Scientific. Room temperature (rt) refers to ambient room temperature (20-22°C), 0°C refers to an ice slush bath and -78°C refers to a dry ice-acetone bath. Heated experiments were conducted using thermostatically controlled oil baths. Reactions were monitored by TLC using aluminium backed silica gel 60 (F254) plates, visualised using UV 254 nm and phosphomolybdic acid or potassium permanganate dips as appropriate. Flash column chromatography was carried out routinely using 60 Å silica gel (Merck). NMR spectra were recorded on a Bruker DPX (300 or 400 MHz) spectrometer. Chemical shifts are reported in δ units, parts per million relative to the singlet at 7.26 ppm for chloroform. Coupling constants (J) are measured in Hertz. IR spectra were recorded on a Perkin-Elmer Spectrum One FT-IR Golden Gate. Mass spectra were recorded on a Bruker Esquire 2000 or a Bruker

MicroTOF mass spectrometer. Melting points were recorded on a Stuart Scientific SMP 1 instrument and are uncorrected. GC analysis was performed using a Hewlett Packard 5890. Optical rotations were measured on an AA-1000 polarimeter. Dry solvents were purchased and used as received. NMR spectra are supplied in Appendix B.

2.1.1 Preparation of 2,3-dihydroxy-2-methylbutanoic acid (11)

Scheme 2.1: Synthetic route to 2,3-dihydroxy-2-methylbutanoic acid (11)



(*E*)-**9** (2*S*,3*R*)-**10** (2*S*,3*R*)-**11**

*2,3-Dihydroxy-2-methylbutanoic acid **11** has been synthesised in an enantiomerically enriched form. The scheme above shows the route to (2S,3R)-**11** but can be used to make all four stereoisomers, using AD-mix α gives (2R,3S)-**11** and using alternative substrate methyl angelate (Z)-**9** and AD-mix β gives (2R,3R)-**11** and with AD-mix α gives (2S,3S)-**11**.*

Four stereoisomers of 2,3-dihydroxy-2-methylbutanoic acid **11** were synthesised in an enantiomerically enriched form in a two-step synthesis that was developed from the protocol described by Stritzke et al. [65], shown in Scheme 2.1. The nomenclature described by Kawai et al. [66] for *threo* and *erythro* diasteromers has been adopted in this work.

2.1.1.1 General procedure for methyl-2,3-dihydroxy-2-methylbutanoate (10)

The relevant AD-mix (11.94 g) was stirred in a 1:1 mixture of tBuOH:H_2O (112 ml) for 5 min at rt. The appropriate alkene starting material **9** (1.0 g, 8.76 mmol) was added dropwise and the solution cooled to 0 °C. The solution was stirred for 24 h and allowed to rise to rt. Sodium

sulphite (12 g) was added and stirred for 90 min. The solution was extracted with EtOAc (4 × 60 ml) and the combined extracts washed with 1 M NaOH (50 ml), brine (50 ml), dried with $MgSO_4$ and solvent removed to give methyl-2,3-dihydroxy-2-methylbutanoate **10** as a clear oil.

2.1.1.1.1 Data for (*threo*)-methyl-2,3-dihydroxy-2-methylbutanoate (*threo*)-10

This compound has been reported and fully characterised [67] and partially characterised [66,68,69].

^{1}H-NMR: δ_H (300 MHz, $CDCl_3$) 3.93 (1H, q, *J* 6.4, CH), 3.79 (3H, s, OCH_3), 1.30 (3H, s, CH_3), 1.20 (3H, d, *J* 6.4, CH_3)

^{13}C-NMR: δ_C (75 MHz, $CDCl_3$) 176.1 (C=O), 76.8 (C), 71.0 (CH), 52.4 (O-CH_3), 21.0 (CH_3), 15.9 (CH_3)

IR: ν_{max} 3454.17 (br, OH), 2984.25 (w, CH), 1731.18 (s, C=O) cm^{-1}

MS (ESI): m/z 171.1 [$(M+Na)^+$]

HRMS (ESI): Measured 171.0628 [$(M+Na)^+$], expected $C_6H_{12}NaO_4$ 171.0628, error 0.0 ppm.

2.1.1.1.2 (2*R*,3*S*)-Methyl-2,3-dihdroxy-2-methylbutanoate (2*R*,3*S*)-10

The general procedure using AD-mix α and methyl tiglate (*E*)-**9** gave (2*R*,3*S*)-methyl-2,3-dihydroxy-2-methylbutanoate (2*R*,3*S*)-**10** (824 mg, 5.57 mmol, 63%).

$[\alpha]_D^{22}$ = +16.4 (c=1.0 in $CHCl_3$, ee 75%), $[\alpha]_D^{20}$ = +21.3 (c=1.1 in $CHCl_3$, ee 77%) (lit[67] $[\alpha]_D^{20}$ = -1.0 (c=1.2, $CDCl_3$)) (lit[69] $[\alpha]_D^{20}$ = +1.0 (c=1, $CHCl_3$)), (lit[70] $[\alpha]_D^{22}$ = -2.5 (neat))

Enantiomeric excess was determined by GC analysis (Chrompac cyclodextrin-β-236M-19 50 m × 0.25 mm × 0.25 μm, Gas: Helium, T = 150°C, P = 15 psi, (2*R*,3*S*) 10.7 min, (2*S*,3*R*) 11.2 min).

2.1.1.1.3 (2*S*,3*R*)-Methyl-2,3-dihydroxy-2-methylbutanoate (2*S*,3*R*)-10

The general procedure using AD-mix β and methyl tiglate (*E*)-**9** gave (2*S*,3*R*)-methyl-2,3-dihydroxy-2-methylbutanoate (2*S*,3*R*)-**10** (737 mg, 4.98 mmol, 57%).

$[\alpha]_D^{22}$ = -17.9 (c=1.0 in $CHCl_3$, ee 87%), $[\alpha]_D^{24}$ = -12.9 (c=1.4 in $CHCl_3$, ee 84%) (lit[68] $[\alpha]_D^{23}$ = +1.0 (c=6.7, $CDCl_3$))

Enantiomeric excess was determined by GC analysis (Chrompac cyclodextrin-β-236M-19 50 m × 0.25 mm × 0.25 μm, Gas: Helium, T = 150°C, P = 15 psi, (2*R*,3*S*) 10.5 min, (2*S*,3*R*) 10.9 min).

2.1.1.1.4 Data for (*erythro*)-methyl-2,3-dihydroxy-2-methylbutanoate (*erythro*)-10

This compound has been reported but it is not fully characterised [66,69,71].

^{1}H-NMR: δ_H (300 MHz, $CDCl_3$) 3.81 (1H, q, *J* 6.5, CH), 3.81 (3H, s, OCH_3), 1.45 (3H, s, CH_3), 1.16 (3H, d, *J* 6.5, CH_3)

^{13}C-NMR: δ_c (75 MHz, $CDCl_3$) 175.3 (C=O), 76.7 (C), 71.6 (CH), 52.2 (O-CH_3), 21.7 (CH_3), 17.1 (CH_3)

IR: ν_{max} 3441.98 (br, OH), 2981.49 (w, CH), 1728.60 (s, C=O) cm^{-1}

MS (ESI): m/z 171.0 [$(M+Na)^+$]

HRMS (ESI): Measured 171.0628 [(M+Na)$^+$], expected $C_6H_{12}NaO_4$ 171.0628 error 0.1 ppm.

2.1.1.1.5 (2*S*,3*S*)-Methyl-2,3-dihydroxy-2-methylbutanoate (2*S*,3*S*)-10

The general procedure using AD-mix α and methyl angelate (*Z*)-**9** was followed except 70 ml of tBuOH:H_2O was used and it was extracted with EtOAc (5 × 50 ml) and washed with 1 M NaOH (10 ml) and brine (25 ml) to give (2*S*,3*S*)-methyl-2,3-dihydroxy-2-methylbutanoate (2*S*,3*S*)-**10** (852 mg, 5.74 mmol, 65%).

$[\alpha]_D^{19}$ = +27.8 (c=1.1 in $CHCl_3$, ee 75%) (lit[71] $[\alpha]_D^{20}$ = +8.7 (c=0.6, $CDCl_3$)) (lit[69] $[\alpha]_D^{20}$ = +12.3 (c=1.0, $CHCl_3$, ee 99.9%))

Enantiomeric excess was determined by GC analysis (Chrompac cyclodextrin-β-236M-19 50 m × 0.25 mm × 0.25 μm, Gas: Hydrogen, T = 95°C, P = 15 psi, (2*R*,3*R*) 28.6 min, (2*S*,3*S*) 29.7 min).

2.1.1.1.6 (2*R*,3*R*)-Methyl-2,3-dihydroxy-2-methylbutanoate (2*R*,3*R*)-10

The general procedure using AD-mix β and methyl angelate (*Z*)-**9** gave (2*R*,3*R*)-methyl-2,3-dihydroxy-2-methylbutanoate (2*R*,3*R*)-**10** (658 mg, 4.45 mmol, 51%).

$[\alpha]_D^{24}$ = -24.8 (c=1.1 in $CHCl_3$, ee 85%)

Enantiomeric excess was determined by GC analysis (Chrompac cyclodextrin-β-236M-19 50 m × 0.25 mm × 0.25 μm, Gas: Hydrogen, T = 95°C, P = 15 psi, (2*R*,3*R*) 29.0 min, (2*S*,3*S*) 31.1 min).

2.1.1.2 General Procedure for 2,3-dihydroxy-2-methylbutanoic acid (11)

The appropriate stereoisomer of methyl-2,3-dihydroxy-2-methylbutanoate **10** (500 mg, 3.38 mmol) was dissolved in iPrOH (30 ml) and added dropwise to a solution of LiOH (450 mg, 18.78 mmol) in water (12 ml) at 0°C with both solutions having been pre-cooled. The reaction was stirred overnight and allowed to rise to rt. 1 M HCl was added until a pH of 2 persisted, solvent was removed to leave ~6 ml, which was extracted with Et_2O (10 $\times$ 30 ml), combined extracts were dried with $MgSO_4$ and solvent removed to leave 2,3-dihydroxy-2-methylbutanoic acid **11**.

2.1.1.2.1 Data for (*threo*)-2,3-dihydroxy-2-methylbutanoic acid (*threo*)-11

This compound has been reported but not fully characterised [72,73].

^{1}H-NMR: δ_H (300 MHz, MeOD) 3.91 (1H, q, *J* 6.4,CH), 1.29 (3H, s, CH_3), 1.18 (3H, d, *J* 6.4, CH_3)

^{13}C-NMR: δ_C (75 MHz, MeOD) 179.0 (C=O), 78.4 (C), 72.7 (CH), 22.3 (CH_3), 16.8 (CH_3)

IR: ν_{max} 3383.57 (br, OH), 2984.74 (w, CH), 1715.34 (s, C=O) cm^{-1}

MS (ESI+): m/z 157.0 [$(M+Na)^+$]

HRMS (ESI): Measured 157.0474 [$(M+Na)^+$], expected $C_5H_{10}NaO_4$ 157.0471 error -1.50 ppm.

2.1.1.2.2 (2*R*,3*S*)-2,3-Dihydroxy-2-methylbutanoic acid (2*R*,3*S*)-11

The general procedure was followed, on a smaller scale. Compound (2*R*,3*S*)-**10** (150 mg, 1.14 mmol) was dissolved in iPrOH (10 ml). LiOH (137 mg, 6.26 mmol) in water (4 ml). After solvent was removed ~2 ml remained, which was extracted with Et_2O (10 $\times$ 10 ml) to give (2*R*,3*S*)-2,3-dihydroxy-2-methylbutanoic acid (2*R*,3*S*)-**11** as a clear oil (103 mg, 0.77 mmol, 67%).

$[\alpha]_D^{29}$ = -0.5 (c=1.05 in MeOH, ee 75%) (lit[74] $[\alpha]_D^{27}$ = +5.7 (c=4.6, H_2O)).

2.1.1.2.3 (2*S*,3*R*)-2,3-Dihydroxy-2-methylbutanoic acid (2*S*,3*R*)-11

The adapted general procedure given for (2*R*,3*S*)-**11** in Section 2.1.1.2.2 was followed using (2*S*,3*R*)-**10** as starting material to give (2*S*,3*R*)-2,3-dihydroxy-2-methylbutanoic acid (2*S*,3*R*)-**11** as a clear oil (77.4 mg, 0.58 mmol, 51%).

$[\alpha]_D^{25}$ = +0.7 (c=1.0 in MeOH, ee 84%).

2.1.1.2.4 Data for (*erythro*)-2,3-dihydroxy-2-methylbutanoic acid (*erythro*)-11

This compound has been reported but not fully characterised [73,75].

^{1}H-NMR: δ_H (300 MHz, MeOD) 3.86 (1H, q, *J* 6.5, CH), 1.40 (3H, s, CH_3), 1.17 (3H, d, *J* 6.5, CH_3)

^{13}C-NMR: δ_c (75 MHz, MeOD) 178.5 (C=O), 78.2 (C), 72.6 (CH), 22.5 (CH_3), 17.7 (CH_3)

IR: ν_{max} 3312.93 (br, OH), 2984.99 (w, CH), 1692.39 (s, C=O) cm^{-1}

MS (ESI): m/z 157.0 $[(M+Na)^+]$

HRMS (ESI): Measured 157.0472 [$(M+Na)^+$], expected $C_5H_{10}NaO_4$ 157.0471 error -0.6 ppm.

2.1.1.2.5 (2*S*,3*S*)-2,3-Dihydroxy-2-methylbutanoic acid (2*S*,3*S*)-11

The general procedure was followed using (2*S*,3*S*)-**10** to give (2*S*,3*S*)-2,3-dihydroxy-2-methylbutanoic acid (2*S*,3*S*)-**11** as a white powder (325 mg, 2.43 mmol, 72%).

$[\alpha]_D^{24}$ = +3.5 (c=1.2 in MeOH, ee 75%) (lit[75] $[\alpha]_D$ = -2.25 (c=3.67, MeOH).

2.1.1.2.6 (2*R*,3*R*)-2,3-Dihydroxy-2-methylbutanoic acid (2*R*,3*R*)-11

The general procedure was followed using (2*R*,3*R*)-**10** to give (2*R*,3*R*)-2,3-dihydroxy-2-methylbutanoic acid (2*R*,3*R*)-**11** as a white powder (276 mg, 2.06 mmol, 61%).

$[\alpha]_D^{25}$ = -4.7 (c=1.0 in MeOH, ee 85%).

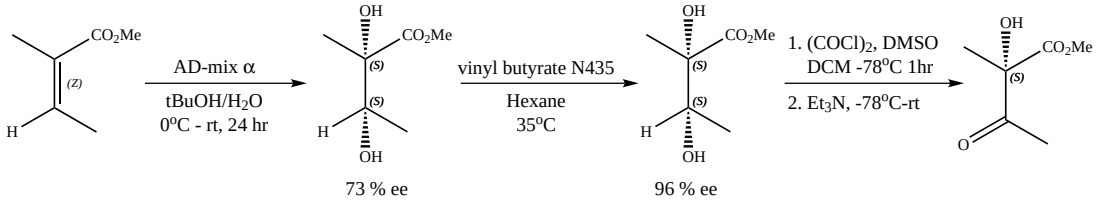
2.1.2 Preparation of (*S*)-methyl acetolactate (*S*)-12

(*S*)-Methyl acetolactate (*S*)-**12** was prepared in a three step synthesis, firstly a dihydroxylation, followed by a kinetic resolution [69] and a Swern oxidation [76].

2.1.2.1 Procedure for (2*S*,3*S*)-methyl-2,3-dihydroxy-2-methylbutanoate (2*S*,3*S*)-10

Compound (2*S*,3*S*)-**10** was synthesised as described in Section 2.1.1.1.5 on a larger scale (methyl angelate (*Z*)-**9** (3.76 g, 32.9 mmol), tBuOH/H_2O (260 ml) and AD-mix α (43.93 g) to give an orange oil as a crude mixture (3.62 g, 73% ee), which was used directly in the next step.

Scheme 2.2: Synthetic route to (*S*)-methyl acetolactate (*S*)-12



73 % ee 96 % ee

(*Z*)-**9** (2*S*,3*S*)-**10** (2*S*,3*S*)-**10** (*S*)-**12**

*(S)-Methyl acetolactate has been synthesised in an enantiomerically enriched form via the three step synthesis shown above. Firstly (Z)-**9** is asymmetrically dihydroxylated to give (2S,3S)-**10** 73% ee. The enantiomeric enrichement is increased by a kinetic resolution using N435 lipase. N435 selectively esterifies the minor isomer, which can then be purified by column chromatography to give (2S,3S)-**10** 96% ee. The final step is a Swern oxidation which oxidises the secondary alcohol to a ketone to give (S)-**12**.*

2.1.2.2 Procedure for kinetic resolution of (2*S*,3*S*)-methyl-2,3-dihydroxy-2-methylbutanoate (2*S*,3*S*)-10

Compound (2*S*,3*S*)-**10** (3.62 g, crude mixture) and vinyl butyrate (454 mg, 3.98 mmol) were dissolved in hexane (70 ml) and stirred at 35°C. The reaction was initiated by addition of the lipase Novozym 435 (900 mg) the reaction was followed by GC and stopped at the desired ee. The desired compound (2*S*,3*S*)-**10** was separated from other compounds by flash chromatography (Hexane:EtOAc). Fractions containing (2*S*,3*S*)-**10** were pooled and solvent removed to give the product (2*S*,3*S*)-**10** as an oil (2.00 g, 13.49 mmol, 41% over 2 steps) The ee was determined by GC to be 96%.

$[\alpha]_D^{30}$ = +12.5 (c=1.2 in $CHCl_3$, ee 96%).

2.1.2.2.1 Analysing the reaction by GC

Reactions were analysed on a Chrompac cyclodextrin-β-236M-19 50 m × 0.25 mm × 0.25 μm column. The column temperature was set at 95°C for 35 min and then ramped at 4°

C/min to 130°C and remained at 130°C for 30 min. Peaks were observed at 29.5 min for (2*R*,3*R*)-methyl-2,3-dihydroxy-2-methylbutanoate (2*R*,3*R*)-**10**, 30.8 min for (2*S*,3*S*)-methyl-2,3-dihydroxy-2-methylbutanoate (2*S*,3*S*)-**10** and 55.3 min for methyl-2-hydroxy-2-methyl-3-(butanoyloxy)butanoate **13**.

2.1.2.2.2 Identification of products on TLC

Three products including the desired compound (2*S*,3*S*)-**10** were separated by TLC, the three compounds were isolated from a silica column and analysed by proton NMR.

The first compound collected was identified by proton NMR as butyric acid **14** [77].

^{1}H-NMR: δ_H (300 MHz, $CDCl_3$) 2.34 (2H, t, *J* 7.3, $CH_2C_2H_5$), 1.67 (2H, sxt, *J* 7.4, CH_2CH_3), 0.98 (3H, t, *J* 7.4, CH_3).

The second compound was identified as methyl-2-hydroxy-2-methyl-3-(butanoyloxy)butanoate **13** by proton NMR [69].

^{1}H-NMR: δ_H (300 MHz, $CDCl_3$) 5.08 (1H, q, *J* 6.5, CH), 3.79 (3H, s, OCH_3), 2.23-51 (2H, m, $CH_2C_2H_5$), 1.49-1.81 (2H, m, CH_2CH_3), 1.37 (3H, s, $H_3CO_2C(OH)CH_3$), 1.17 (3H, d, *J* 6.4, $CHCH_3$) 0.88-0.99 (3H, m, $C_2H_4CH_3$).

The third product was identified as the desired compound (2*S*,3*S*)-methyl-2,3-dihydroxy-2-methylbutanoate (2*S*,3*S*)-**10** (for full data see Section 2.1.1.1.4).

^{1}H-NMR: δ_H (300 MHz, $CDCl_3$) 3.81 (1H, q, *J* 6.5, CH), 3.81 (3H, s, OCH_3), 1.45 (3H, s, CH_3), 1.16 (3H, d, *J* 6.5, CH_3).

Enantiomeric excess was determined by GC analysis (Chrompac cyclodextrin-β-236M-19 50 m × 0.25 mm × 0.25 μm, Gas: Hydrogen, T = 95°C, P = 15 psi, (2*R*,3*R*) 28.8 min,

(2*S*,3*S*) 29.6 min).

2.1.2.3 Procedure for the synthesis of (*S*)-methyl acetolactate (*S*)-12

This compound has been reported and fully characterised [78,79], the enantiomer (*R*)-methyl acetolactate (*R*)-**12** has been reported and characterised [55] and racemic methyl acetolactate **12** has been reported and characterised [80].

Oxalyl chloride (2 M in DCM, 3.35 ml) and DMSO (0.95 ml, 13.4 mmol) were stirred in DCM (6.7 ml) under inert conditions at -78°C for 15 min. Then substrate (2*S*,3*S*)-**10** (979 mg, 6.61 mmol) in DCM (1 ml) was added dropwise and stirred for 60 min at -78°C. Et_3N (4.9 ml, 35 mmol) was added and allowed to warm up to rt. The compound was purified by flash chromatography using a mixture of pentane:diethyl ether, fractions containing the compound were pooled and solvent removed under low vacuum to give a yellow oil (668 mg, 4.57 mmol, 69%).

^{1}H-NMR: δ_H (300 MHz, $CDCl_3$) 4.22 (1H, s, OH), 3.81 (3H, s, CH_3), 2.29 (3H, s, CH_3), 1.61 (3H, s, CH_3)

^{13}C-NMR: δ_c (176 MHz, $CDCl_3$) 204.9 (C=O), 171.8 (C=O), 81.0 (C), 53.4 (OCH_3), 24.2 (CH_3), 22.0 (CH_3)

IR: νmax 3460 (br, OH), 2958 (w, CH), 1717 (s, C=O) cm^{-1}

MS (ESI): m/z 169.1 [$(M+Na)^+$]

HRMS (ESI): Measured 169.0470 [$(M+Na)^+$], expected $C_6H_{10}NaO_4$ 169.0471 error 0.6 ppm.

$[\alpha]_D^{28}$ = +13.2 (c=1.3 in EtOH, ee 96%)(lit[79] $[\alpha]_D$ = +9.7 (c=3.1 in EtOH))(lit[55] (*R*)-methyl acetolactate $[\alpha]_D^{30}$ = -9.6 (c=3.1 in EtOH) $[\alpha]_D^{30}$ = +27.3 (c=6.4 in H_2O)).

2.1.2.3.1 Identification of Swern oxidation side product

In the Swern oxidation protocol that was followed there was a footnote describing a methylthiomethylation side product. This compound was eluted from a silica gel column earlier than the desired (*S*)-methyl acetolactate (*S*)-**12**. The methylthiomethylation product (*S*)-methyl-2-methyl-2-(methylsulphanylmethoxy)-3-oxobutanoate (*S*)-**15** was isolated and analysed, as it is not reported and there is no data available on this compound in the literature.

^{1}H-NMR: δ_H (300 MHz, $CDCl_3$) 4.69 (2H, s, CH_2), 3.77 (3H, s, OCH_3), 2.30 (3H, s, CH_3), 2.24 (3H, s, SCH_3), 1.60 (3H, s, $CO(CH_3)$)

^{13}C-NMR: δ_c (75 MHz, $CDCl_3$) 204.6 (C=O), 170.0 (C=O), 85.9 (C), 70.2 (CH_2), 52.7 (OCH_3), 25.2 (CH_3), 18.0 (CH_3), 14.6 (SCH_3)

IR: ν_{max} :2924 (w, CH), 1747 (s, C=O), 1720 (s, C=O) cm^{-1}

MS (ESI): m/z 229.0 $[(M+Na)^+]$

HRMS (ESI): Measured 229.0504 $[(M+Na)^+]$, expected $C_8H_{14}NaO_4S$ 229.0505 error 0.3 ppm.

$[\alpha]_D^{30}$ = +27.5 (c=0.7 in $CHCl_3$, ee 96%).

2.1.2.4 Preparation of (*S*)-acetolactate (*S*)-3

(*S*)-Acetolactate (*S*)-**3** is the substrate for ALDC and is unstable so had to be prepared by hydrolysis of (*S*)-methyl acetolactate (*S*)-**12** immediately before use.

2.1.2.4.1 Procedure for (*S*)-acetolactate (*S*)-3

(*S*)-Methyl acetolactate (*S*)-**12** (4.0 μl, 27.4 nmol) was stirred in 0.5 M NaOH (150 μl) for 5 min at rt. Then 0.1 M phosphate buffer pH 6.0 (250 μl) was added and HCl added to give a pH of 6.0 and diluted with buffer to give the desired concentration.

2.1.2.4.2 Confirmation of hydrolysis by CD

The CD spectrum of (*S*)-methyl acetolactate (*S*)-**12** (1% v/v, 95% ee) was measured on a J-715 spectropolarimeter (Jasco, UK) using a 1 mm path length cuvette. Circular dichroism spectra were recorded between 180 and 450 nm with a data pitch of 0.2 nm, a bandwidth of 2 nm, a scanning speed of 100 nm/min and a response time of 1 sec. To monitor the hydrolysis, 0.5 M NaOH (120 μl) was added and the spectrum was recorded. The spectrum of (*S*)-acetolactate (*S*)-**3** agreed with the literature and showed that the reaction had gone to completion in 2 min.

2.1.2.4.3 Assessing the stability of (*S*)-acetolactate (*S*)-3 at different pH

The stability of (*S*)-acetolactate (*S*)-**3** was monitored at pH 14, 12.7 and 7. (*S*)-Methyl acetolactate (*S*)-**12** (2 μl) was dissolved in water (198 μl), 0.5 M NaOH (120 μl) was added and stirred for 5 min. The pH 14 sample was diluted to 1 ml with 0.5 M NaOH, the pH 12.7 sample was diluted to 1 ml with water and the pH 7 sample was diluted with 0.1 M phosphate buffer pH 7 and acidified with HCl until the pH measured pH 7 and a total volume of 1 ml. The sample (200 μl) was placed in a 1 mm quartz cuvette and CD spectra recorded between 190 and 400 nm with a data pitch of 0.2 nm, a bandwidth of 2 nm, a scanning speed of 100 nm/min and a response time of 2 min. At pH 14 scans were performed every 10 min for 40

scans, at pH 12.7 scans were performed every 30 min for 40 scans and at pH 7 scans were performed every 30 min for 32 scans.

2.2 Cloning, expression and purification of *Bacillus subtilis* acetolactate decarboxylase and acetoin reductase

2.2.1 Growth and maintenance of *E.coli* strains

2.2.1.1 Bacterial growth media

Luria Bertani (LB) broth, Super Optimal broth with Catabolite repression (SOC) and LB agar plates were prepared by the Media Preparation Service at the University of Warwick.

2.2.1.2 Bacterial strains

Table 2.1: Bacterial strains used in this study

***E. coli* strain**	**Genotype**
TOP10	F^- *mcrA* Δ*(mrr-hsdRMS-mcrBC)* φ80*lacZ*ΔM15 Δ*lacX*74 *deoR recA*1 *araD*139 Δ (*ara-leu*)7697 *galU galK rpsL* (Str^R) *endA1 nupG*
B834(DE3) pRare	F^- *ompT hsd*S_B(r_B^- m_B^-) *gal dcm met* (DE3) pRARE (Cam^R)

2.2.1.3 DNA transformation of competent *E. coli*

A 50 µl aliquot of competent cells was incubated on ice with 1 µl of DNA for 30 min. They were heatshocked by incubating at 42°C for 30 sec and incubating on ice for 2 min. SOC

(300 μl) was added to the cells, which were incubated for 1 hour at 37°C with shaking. The cells were spread onto agar plates containing the appropriate antibiotic for selection and incubated overnight at 37°C.

2.2.2 DNA manipulation and cloning techniques

2.2.2.1 Primers

Primers used in this study were designed according to the DNA manipulation technique required and ordered from Integrated DNA Technologies (UK).

Table 2.2: Primers used for generation of constructs

Primer name	Restriction site	Sequence 5' to 3'
bdhA_for	*Bam*HI	C TTT TGC **GGA TCC** ATG AAG GCA GCA AGA TGG CAT
bdhA_rev	*Xho*I	TTT GCG **CTC GAG** TTA GTT AGG TCT AAC AAG GAT
alsD_for	*Bam*HI	C TTT TGC **GGA TCC** ATG AAA CGA GAA AGC AAC ATT
alsD_rev	*Xho*I	TTT GCG **CTC GAG** TTA TTC AGG GCT TCC TTC AGT

Primers used for generation of constructs are shown in the 5'-3' direction, the forward primers contain a BamHI restriction site shown in bold and the reverse primers contain a XhoI restriction site shown in bold.

2.2.2.2 PCR

PCR was typically conducted, according to the manufacturer's instructions, with 5 μl of 10 × buffer, 1.5 μl of 10 mM deoxynucleotide triphosphates (dNTPs), 1μl of 50 mM of $MgSO_4$, 5 μl of 10 × enhancer, 1.5 μl of forward primer, 1.5 μl of reverse primer, 1 μl of template DNA, 0.5 μl of Pfx in a total volume of 50 μl.

Reactions were performed on an Eppendorf Mastercycler Gradient Thermocycler using the programme shown in Figure 2.1.

Figure 2.1: PCR protocol

94℃	2 min	
94℃	30 sec	$\times 40$
55℃	30 sec	
68℃	2 min	
68℃	10 min	
4℃	hold	

2.2.2.3 Restriction endonuclease digestion of DNA

DNA was digested with BamHI (1 μl) and XhoI (2 μl) (Invitrogen) in Buffer React3 in a total volume of 20 μl for 4 hours at 37°C.

2.2.2.4 Purification of DNA from PCR reaction and restriction digests

DNA was purified directly from the PCR or restriction digest mix to remove polymerases or endonucleases using a QIAquick Gel Extraction Kit (Qiagen, Germany) according to the manufacturer's instructions.

2.2.2.5 Agarose gel electrophoresis

Agarose gels were prepared by dissolving 0.8% (w/v) agarose in TAE buffer (40 mM Tris acetate, 1 mM EDTA). Ethidium bromide was added to allow visualisation of the DNA. DNA samples were prepared by adding 6 × loading dye (Fermentas, 10 mM Tris-HCl pH 7.6, 0.03% bromophenol blue, 0.03% xylene cyanol FF, 60% glycerol, 60 mM EDTA). Gels were

placed in a gel tank and submerged with TAE buffer, samples were loaded and 5 µl of 1 kb DNA ladder (Fermentas) was added to allow size determination. Electrophoresis was performed at 150 V for 30 min and the gel was visualised with ultraviolet light using a Syngene GeneSnap G:box illuminator and analysis system.

2.2.2.6 Purification of digested vector DNA from agarose gel

The band containing the digested vector DNA was excised from the gel and purified using a QIAquick Gel Extraction Kit (Qiagen, Germany) according to the manufacturer's instructions.

2.2.2.7 Ligations

Ligation reactions using T4 DNA ligase (Invitrogen) were set up according to the manufacturer's instructions. Vector and insert fragments were incubated in 1:1, 1:3 and 1:10 molar ratios and left at rt overnight.

2.2.2.8 Preparation of plasmid DNA

Plasmid DNA was isolated using a QIAprep Mini-Prep kit (Qiagen, Germany). The kit was used according to the manufacturer's instructions, plasmid DNA was stored at -20°C.

2.2.2.9 DNA concentration determination

DNA quantification was determined using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific) using 1.5 µl samples.

2.2.2.10 Sequencing of plasmid DNA

Plasmid DNA was sequenced to confirm generation of the construct. Plasmid DNA (250-400 ng) and T7 sequencing primers (10 pmol) were diluted to a volume of 10 μl with water and were submitted to the Molecular Biology Service at the University of Warwick.

2.2.2.11 Site-directed mutagenesis

A series of point mutations were introduced into *Bacillus subtilis* ALDC using the QuickChange Lightning kit (Agilent) according to the manufacturer's instructions. Primers shown in Table 2.3 were designed according to the manufacturer's instructions, the substituted codons are underlined and the mutated bases shown in bold.

Table 2.3: Primers used for site-directed mutagenesis

Mutation	Primer name	Sequence 5' to 3'
T55S	T55S_for	A GAC TTC GGT ATC GGA A**G**C TTT AAC AAG CTT GAC G
	T55S_rev	C GTC AAG CTT GTT AAA G**C**T TCC GAT ACC GAA GTC T
T55A	T55A_for	GGA GAC TTC GGT ATC GGA **G**CC TTT AAC AAG C
	T55A_rev	G CTT GTT AAA GG**C** TCC GAT ACC GAA GTC TCC
E62Q	E62Q_for	GAA CCT TTA ACA AGC TTG ACG GA**C** AGC TGA TTG GGT
	E62Q_rev	ACC CAA TCA GCT **G**TC CGT CAA GCT TGT TAA AGG TTC
E62A	E62A_for	T AAC AAG CTT GAC GGA G**C**G CTG ATT GGG TTT GAC G
	E62A_rev	C GTC AAA CCC AAT CAG C**G**C TCC GTC AAG CTT GTT A
R142K	R142K_for	GAC GGA TTG TTT AAA AAG GTG CAG ACA A**A**A ACA GTA GAA CTT CAA G
	R142K_rev	C TTG AAG TTC TAC TGT T**T**T TGT CTG CAC CTT TTT AAA CAA TCC GTC
R142A	R142A_for	GGA TTG TTT AAA AAG GTG CAG ACA **GC**A ACA GTA GAA CTT CAA GAA AAA CC
	R142K_rev	GG TTT TTC TTG AAG TTC TAC TGT T**GC** TGT CTG CAC CTT TTT AAA CAA TCC
E251Q	E251Q_for	CT GAT TTT GCG AAA GAT ATC GAA ACA ACT **CAG** GGA AGC CCT GAA TAA
	E251Q_rev	TTA TTC AGG GCT TCC **CTG** AGT TGT TTC GAT ATC TTT CGC AAA ATC AG
E251A	E251A_for	G AAA GAT ATC GAA ACA ACT G**C**A GGA AGC CCT GAA TAA CTC G
	E251A_rev	C GAG TTA TTC AGG GCT TCC T**G**C AGT TGT TTC GAT ATC TTT C

The primers used for site-directed mutagenesis are shown in the 5'-3' direction, the codons for the effected amino acids are underlined with the mutated bases shown in bold.

2.2.2.11.1 Plasmids

Table 2.4: Commercial plasmids used in this study

Plasmid	Details	Supplier
pET 28a	Cloning vector	Novagen

Table 2.5: Plasmids generated in this study

Plasmid	Template plasmid	Primers used	Features
pET 28 bdhA	pET 28a	bdhA_for bdhA_rev	*B. subtilis* Acetoin reductase with a N-terminal hexa-histidine tag
pET 28 alsD	pET 28a	alsD_for alsD_rev	*B. subtilis* Acetolactate decarboxylase with a N-terminal hexa-histidine tag
pET 28 alsD T55S	pET 28 alsD	T55S_for T55S_rev	*B. subtilis* Acetolactate decarboxylase Thr 55 to Ser mutation
pET 28 alsD T55A	pET 28 alsD	T55A_for T55A_rev	*B. subtilis* Acetolactate decarboxylase Thr 55 to Ala mutation
pET 28 alsD E62Q	pET 28 alsD	E62Q_for E62Q_rev	*B. subtilis* Acetolactate decarboxylase Glu 62 to Gln mutation
pET 28 alsD E62A	pET 28 alsD	E62A_for E62A_rev	*B. subtilis* Acetolactate decarboxylase Glu 62 to Ala mutation
pET 28 alsD R142K	pET 28 alsD	R124K_for R142K_rev	*B. subtilis* Acetolactate decarboxylase Arg 142 to Lys mutation
pET 28 alsD R142A	pET 28 alsD	R142A_for R142A_rev	*B. subtilis* Acetolactate decarboxylase Arg 142 to Ala mutation
pET 28 alsD E251Q	pET 28 alsD	E251Q_for E251Q_rev	*B. subtilis* Acetolactate decarboxylase Glu 281 to Gln mutation
pET 28 alsD E251A	pET 28 alsD	E251A_for E251A_rev	*B. subtilis* Acetolactate decarboxylase Glu 281 to Ala mutation

2.2.3 Protein expression and purification

2.2.3.1 Over-expression in E. coli

Recombinant proteins were over-expressed in *E. coli* B834(DE3), plasmids were transformed into competent *E. coli* B834(DE3) cells and plated onto agar plates containing the appropriate antibiotic. A single colony was used to inoculate a small scale culture of 5 ml LB containing kanamycin (50 μg/ml) and chloramphenicol (35 μg/ml). The culture was incubated at 37°C overnight with shaking at 180 rpm. The entire 5 ml culture was used to inoculate a 1 l expression culture in the same medium, which was grown at 37°C with shaking at 180 rpm until an OD_{600} of 0.6-0.8 was reached. Cells were induced with 0.5 mM IPTG for 3-4 h at 37°C with shaking at 180 rpm. Cells were harvested by centrifugation at 6000 rpm for 10 min in a Beckman JLA-8.100 rotor at 4°C. Cell pellets were stored at -80°C.

2.2.3.2 Preparation of crude cell lysate

Cell pellets were thawed and resuspended in buffer A (20 mM HEPES pH 8.0, 0.5 M NaCl, 10 mM imidazole, 10% (v/v) glycerol). Cells were lysed by 3 $\times$ 30 sec blasts of sonication using a Bandelin Sonoplus sonicator at 70% power. The cell debris was removed by centrifugation at 20,000 rpm for 30 min at 4°C in a Beckman JA-25.50 rotor. The supernatant containing the soluble fraction was retained for purification.

2.2.3.3 Protein purification

2.2.3.3.1 Immobilised metal affinity chromatography

As the proteins contain an N-terminal hexa-histidine tag, the proteins were purified on a 5 ml HisTrap (GE Healthcare) Column using an ÄKTApurifier. Supernatant was loaded onto the column that had been pre-equilibrated with buffer A (20 mM HEPES pH 8.0, 0.5 M NaCl, 10 mM imidazole, 10% glycerol). The column was washed with buffer A until the UV 254 nm and 280 nm had stabilised to remove unbound proteins. The column was washed with 10% buffer B (20 mM HEPES pH 8.0, 0.5 M NaCl, 500 mM imidazole) followed by 20% buffer B to remove loosely associated proteins. Finally the protein was eluted with 100% buffer B, 10 ml fractions were collected. Fractions that contained protein were pooled and buffer exchanged. All AlsD samples were exchanged into a 20 mM Tris pH 8.0 buffer by concentrating with a centrifugal concentrator 20000 MW cut off and diluting with buffer and repeating until the concentration of imidazole was diluted to less than 10 mM. BdhA was exchanged into a buffer containing 20 mM Tris pH 8.0 and 100mM NaCl by dialysis and concentrated with a 20000 MW cut off centrifugal concentrator.

2.2.3.3.2 Size exclusion chromatography

Protein in gel filtration buffer (50 mM Tris pH 8.0, 150 mM NaCl) was loaded via a 0.5 ml injection loop onto a Superdex S200 (GE healthcare) column that had been pre-equilibrated in gel filtration buffer. Buffer was washed through at a flow rate of 1 ml/min, 10 ml fractions were collected and protein elution was monitored by absorbance at A_{280}.

2.2.3.3.3 Anion exchange chromatography

Protein was loaded onto a Resource Q column (GE healthcare) in a low salt buffer (50 mM Tris pH 8.0). After the column had equilibrated a salt gradient was applied to the column from 0-1 M NaCl (in 50 mM Tris pH 8.0) over 20 column volumes at a flow rate of 1 ml/min. Fractions of 1 ml were collected and protein elution was followed by absorbance at A_{280}.

2.2.3.3.4 Hydrophobic interaction chromatography

Protein was loaded onto a Resource HIC column (GE healthcare) in a high salt buffer (50 mM Tris pH 8.0, 1 M $(NH_4)_2SO_4$). After the column had equilibrated a salt gradient was applied to the column from 1-0 M $(NH_4)_2SO_4$) (in 50 mM Tris pH 8.0) over 20 column volumes at a flow rate of 1 ml/min. 1 ml fractions were collected and protein elution was followed by absorbance at A_{280}.

2.2.4 Protein analysis and detection

2.2.4.1 SDS polyacrylamide gel electrophoresis

Sodium dodecyl sulphate (SDS)-polyacrylamide gels were cast and run on a vertical gel electrophoresis system (Hoefer Mighty Small) based on the method described by Laemmli [81]. Typically 1 mm gels were prepared with a 14% resolving gel (375 mM Tris pH 8.8, 14% acrylamide:bis-acrylamide (29:1), 0.1% SDS, 0.1% ammonium persulphate (APS), 0.1% Tetramethylethylenediamine (TEMED) and a stacking gel (125 mM Tris pH 6.8, 4% acrylamide:bis-acrylamide (29:1), 0.1% SDS, 0.1% APS, 0.2% TEMED).

Samples were prepared by mixing with 5 $\times$ SDS loading buffer (100 mM Tris pH 6.8, 2%

(w/v) SDS, 15% (v/v) glycerol, 5% β-mercaptoethanol, 0.02% bromophenol blue) and heating to 95°C for 2 min in an Eppendorf Mastercycler Gradient Thermocycler.

Gels were run in SDS running buffer (25 mM Tris pH 8.3, 0.19 M glycine, 0.1% (w/v) SDS) at 180 V until the dye front had run off the end of the gel (typically 55 min).

2.2.4.2 Development of SDS PAGE

Proteins separated by SDS-PAGE were detected using Colloidal-Coomassie stain (8% (w/v) ammonium sulphate, 0.08% (w/v) Coomassie G-250, 1.6% (v/v) phosphoric acid, 20% (v/v) ethanol) or Instant Blue (Expedeon). Following SDS-PAGE, when using Colloidal-Coomassie gels were rinsed in water for 5 min before staining overnight and destained in water. When using instant blue, gels were stained for up to 1 h. Gels were imaged using a Syngene GeneSnap G:Box Gel Doc and analysis system.

2.2.4.3 Western blot

Polypeptides that had been resolved by SDS-PAGE were transferred from the gel to a polyvinylidene difluoride (PVDF) membrane. The transfer took place in Western Transfer buffer (50 mM Tris, 40 mM Glycine, 0.04% (w/v) SDS, 20% (v/v) MeOH) at 100 V for 1 h using an Electrophoretic Protein Blotting Tank (BioRad). The PVDF membrane was washed with PBST (1 X PBS, 1% (v/v) Tween-20) and blocked with PBST containing 5% (w/v) skimmed milk powder. The membrane was washed with PBST and incubated overnight in 20 ml PBST containing 10 µl anti-histidine tagged mouse monoclonal antibody (Merck-Millipore, clone 4D11, 05-531) at 4°C. The membrane was washed with PBST three times for 5 min. The blot was incubated in 20 ml PBST containing 2 µl of the secondary goat antibody raised against anti-mouse IgG conjugated to horse-radish peroxidase (Novagen) for 1 h at 4°C

and 1 h at rt with shaking. The blot was washed three times for 5 min with PBST and reacted with developer solution (ECL Plus Western Blotting Detection System, Amersham GE Healthcare), which was prepared according to the manufacturers instructions for 3 min. The blot was placed in a cassette containing KODAK BioMax MR film and exposed for 1-5 min. The film was developed using an AGFA Curix 60 developing machine.

2.2.4.4 Mass spectrometry

2.2.4.4.1 Gel based protein identification

Coomassie-stained SDS PAGE gels were submitted to the Warwick/Waters Centre for BioMedical Mass Spectrometry and Proteomics at the University of Warwick for tryptic digest and protein identification by nanoLC-ESI-MS/MS.

2.2.4.4.2 Intact molecular mass analysis

Concentrated purified proteins were submitted to the Warwick/Waters Centre for Biomedical Mass Spectrometry and Proteomics at the University of Warwick for intact molecular mass analysis by positive ion ESI-MS.

2.2.4.5 Circular dichroism

Protein samples were prepared in 10 mM sodium phosphate buffer pH 7.6 to give a final concentration of 0.1 mg/ml. Measurements were recorded on a J-815 spectropolarimeter (Jasco,UK) using a 1 mm path length cuvette. Circular dichroism spectra were recorded between 180 and 260 nm with a data pitch of 0.5 nm, a bandwidth of 1 nm, a scanning

speed of 100 nm/sec and a response time of 1 sec. Nine scans were averaged and the spectrum of buffer subtracted to give the final spectra.

2.2.4.5.1 Secondary structure prediction

CD spectra were analysed using the CDSSTR programme of Dichroweb using reference set SP175 Optimised for 190 - 240 nm [82].

2.2.5 Homology modelling

Bacillus subtilis ALDC sequence was submitted to Phyre2 a web based protein structure prediction tool [83].

2.3 Assay of ALDC activity

2.3.1 Discontinuous Voges-Proskauer assay

Reagents were purchased from Sigma-Aldrich and used according to the instructions provided. The Sigma-Aldrich protocol is supplied in Appendix C.

2.3.1.1 Preparation of (±)-acetolactate (3)

Substrate solution (10 mM stock) was prepared by hydrolysing 10 µl of ethyl-2-acetoxy-2-methylacetoacetate by stirring for 20 min in the presence of 600 µl of 0.5 M NaOH. The solution was then adjusted to ~4 ml by addition of buffer A (50 mM MES pH 6.0) and the pH adjusted to 6.0 using 1 M HCl, then diluted with buffer A to a final volume of 5 ml. Substrate

solution was diluted with buffer A (50 mM MES pH 6.0) to give the following concentrations; 0.1 mM, 0.5 mM, 1 mM, 2.5 mM, 3.75 mM, 5 mM, and 7.5 mM.

2.3.1.2 Preparation of *B. brevis* ALDC solution

Enzyme solution was prepared by diluting 20 μl of ALDC (11 μg/ml) with buffer C (50 mM MES Buffer pH 6.0 with 0.005% Brij 35 Solution and 600 mM NaCl) to give a final volume of 4 ml.

2.3.1.3 Stopped assay and product detection

Substrate (200 μl) was incubated with enzyme solution (200 μl) for 20 min at 30 °C. For each sample, a blank containing no enzyme solution was treated the same way. During sample incubation, colour reagent (100 ml) was prepared by dissolving α-naphthol (1 g) and creatine (0.1 g) into a minimum amount of 1 M NaOH, made up to 100 ml and protected from the light until use. Samples were removed from the water bath and incubated with 4.5 ml of colour reagent (for the blanks, enzyme solution was also added) and incubated for 40 min at rt for colour development. The absorbance at 522 nm was then measured.

2.3.1.4 Standard curve

A standard curve was produced using a stock acetoin solution. Racemic acetoin (±)-**4** was dissolved in deionised water to give a 2.27 mM stock solution. The stock solution was diluted with water to give a final volume of 400 μl containing 11.35 μM, 22.7 μM, 45.4 μM, 68.1 μM, 90.8 μM of acetoin. Colour reagent (4.5 ml) was added to each sample and to the blank that contained just water and colour reagent. The colour was allowed to develop over 1 h at rt and the absorbance was recorded at 522 nm.

2.3.2 Continuous spectrophotometric coupled assay

This assay was developed to detect the activity of purified ALDC by following the disappearance of NADH. The coupling enzyme acetoin reductase converts the acetoin produced by ALDC to butane-2,3-diol with the oxidation of NADH to NAD^+, which can be followed spectrophotometrically at 340 nm.

2.3.2.1 Assay to detect coupling enzyme activity

The activity of the coupling enzyme was determined according to the method described by Yan et al. [53]. Assay reactions contained 50 mM Tris-HCl pH 8.0, 100 μM NADH and 0.01-6 mM acetoin to give a total volume of 1 ml. Assays were conducted on a Varian Cary spectrophotometer at 20°C, reactions were initiated by the addition of BdhA to give a final concentration of 9 μM.

2.3.2.2 Coupled assay

A 10 mM stock of racemic acetolactate (±)-**3** was prepared as described in 2.3.1.1. Assays were conduced on a Varian Cary spectrometer at 20°C. The following were mixed together in a cuvette to give a final volume of 1 ml: 100 mM Bis-tris pH 7; 10 mM NADH; 0-2.5 mM acetolactate **3**; BdhA. Reactions were initiated by the addition of ALDC. The disappearance of NADH was monitored at 340 nm and the gradient was measured whilst the rate was linear.

2.3.3 Continuous circular dichroism assay

Assays were conducted on a Jasco J-720 Circular Dichroism spectrophotometer with Peltier temperature control unit. Substrate (*S*)-acetolactate (*S*)-**3** was prepared immediately before

use (Section 2.1.2.4.1). Samples were prepared by mixing (*S*)-**3** (0.5-30 mM) with 0.1 M phosphate buffer pH 7 to a final volume of 200 μl. Using a 1 mm quartz cuvette, the change in CD at 278 nm was observed. Reactions were initiated by the addition of enzyme.

2.3.3.1 Mutant activity assays

Initial mutant activities were tested at a single substrate concentration. Samples were made up to 200 μl containing 5 mM (*S*)-**3** in 0.1 M phosphate buffer pH 7, reactions were performed in triplicate and initiated by the addition of 2 μl of enzyme. The active mutants were fully characterised as described for wild type ALDC in Section 2.3.3.

2.3.3.2 Inhibition assays

Samples were made up to 200 μl containing (*S*)-**3** (0.5, 1.0, 2.5, 5.0, 10.0 mM), inhibitor (0, 1.0, 2.5, 5.0, 10.0, 15.0 mM) and 0.1 M phosphate buffer pH 6. Reactions were initiated by the addition of enzyme to give a final concentration of 9.12 nM.

2.3.4 Molar circular dichroism

To convert rates measured in mdeg/s into meaningful rates of enzyme kinetics, the molar circular dichroism ($\Delta\varepsilon$) was measured.

2.3.4.1 (*S*)-acetolactate (*S*)-3 sample preparation

(*S*)-Methyl acetolactate (*S*)-**12** (2.92 μl, 20 nmol, 95% ee) was stirred in 0.5 M NaOH (240 μl) for 5 min at rt. The reaction was diluted with 0.1 M sodium phosphate buffer pH 6.5 (500 μl) and HCl added until the pH measured 6.5. Sodium phosphate buffer was added to give

a total volume of 1 ml and a stock concentration of 20 mM (*S*)-acetolactate (*S*)-**3**. This was diluted to give the following concentrations: 15 mM, 10 mM, 5 mM, 1 mM and 0.5 mM.

2.3.4.2 (*R*)-acetoin (*R*)-4 sample preparation

A sample of 20 mM (*S*)-acetolactate (500 μl) (*S*)-**3** prepared in 2.3.4.1 was transferred to a fresh vial and stirred at rt in the presence of *B. brevis* ALDC (0.73 μM) for 30 min. To give a stock solution of 20 mM (*R*)-acetoin (*R*)-**4**. This was diluted to give the following concentrations: 15 mM, 10 mM, 5 mM, 1 mM and 0.5 mM.

2.3.4.3 Sample measurement

The 10 samples described in Sections 2.3.4.1 and 2.3.4.2 were transferred to a 1 mm quartz cuvette and the CD spectrum from 190-340 nm was measured with 5 accumulations. The baseline spectrum of buffer was subtracted to give the final spectrum.

The molar circular dichroism of (*S*)-acetolactate (*S*)-**3** was measured at 312 nm and (*R*)-acetoin (*R*)-**4** at 278 nm. The value was calculated from the gradient of a plot of concentration versus molar circular dichrosim.

2.4 Crystallography of *Bacillus brevis* acetolactate decarboxylase

ALDC from *Bacillus brevis* was supplied as a gift from Novozymes. It was provided as a 3.15 mg/ml solution in 50 mM Tris-HCl pH 8.5 and stored at -80°C. For crystallisation studies ALDC was concentrated to 13.5 mg/ml (50 mM Tris-HCl pH 8.5) using a Vivaspin centrifugal

concentrator with a molecular weight cut off of 10,000 Da.

2.4.1 Replicating previous crystallisation conditions

ALDC had previously been crystallised under three different conditions to give rise to three different crystals forms [63].

2.4.1.1 Type 1 crystals

Type 1 crystals were reproduced under the following conditions; 15-20% PEG 8K, 0.1 M MES pH 6.5 and 200 mM $Zn(OAc)_2$. Crystals were grown in 24 well plates with a reservoir volume of 200 μl and hanging drops on plastic coverslips containing 2 μl of ALDC (11 mg/ml) mixed with 1 μl of mother liquor.

2.4.1.2 Type 2 crystals

Type 2 crystals were irreproducible under the conditions tried. Conditions were screened around 12.5-30% PEG 2K MME, 0.1 M Tris-HCl pH 6.0-9.0 and 5-40 mM $CdCl_2$. Needle clusters were observed but the rectangular crystals previously described were not observed.

2.4.2 Commercial screens

As attempts to reproduce the high resolution type 2 crystals were unsuccessful commercial screens were used to find new conditions.

Initial crystallisation screens were performed using JCSG-*plus* and PACT *premier* 96 well screens from Molecular Dimensions on the Honeybee 963 crystallisation robot (Digilab) at

18°C. Each well contained 70 μl precipitant and sitting drops contained 200 nl of precipitant mixed with 200 nl *B. brevis* ALDC (a gift from Novozymes) at 13.5 mg/ml.

2.4.3 Optimisation

Crystals appeared in well G4 (0.2 M TMAO, 0.1 M Tris pH 8.0, 20% PEG 2K MME) of the JCSG+ screen after 3 months. The conditions were optimised in hanging drop 24 well plates. Crystalline material appeared after 3 months and was used as microseeds to obtain single crystals, using 0.2 M TMAO, 0.1 M Tris pH 8.5, 28-33% PEG 2K MME.

2.4.4 Co-crystallisation

Protein-ligand complex structures were obtained by growing crystals in the presence of 2,3-dihydroxy-2-methylbutanoic acid **11**. 1 M inhibitor (suspended in water) was added to the protein in a 1:9 ratio and trays set up as described in Section 2.4.3

2.4.5 Cryoprotection

2.4.5.1 Cryoprotection with LV CryoOil

Crystals were cryoprotected with LV CryoOil (MiTeGen), a single crystal was transferred on a nylon loop into a 1 μl droplet of oil and held there for 30 sec. For in-house X-ray diffraction data collection, the crystal was transferred directly into the cryostream, whereas for synchrotron X-ray diffraction data collection crystals were cryo-cooled in liquid nitrogen.

2.4.5.2 Cryoprotection with glycerol

Crystals were transferred on a nylon loop to a solution containing 30% PEG 2K MME, 25% (v/v) glycerol, 0.1 M TMAO and 0.05 M Tris-HCl pH 8.0 and held there for 30 sec before being transferred directly into the cryostream.

2.4.6 Data collection

Data were collected in-house on a mar345 image plate detector or at beamline IO4 at the Diamond Light Source. Table 2.6 outlines the crystal structures collected during this work, the cryoprotection method and the data collection source.

Table 2.6: Crystal structures obtained during this work

Crystal structure name	Cryoprotection method	Data collection source
ALDC-PO4	LV CryoOil	DLS Beamline I04
ALDC-GOL	Glycerol	mar345 image plate detector
ALDC-SS	LV CryoOil	mar345 image plate detector
ALDC-RR	LV CryoOil	DLS Beamline I04
ALDC-SR	LV CryoOil	DLS Beamline I04

2.4.6.1 Home source

ALDC-SS and ALDC-GOL datasets were collected on a mar345 image plate detector in-house. ALDC-SS was exposed for 120 sec per image, 360° of data were collected with a rotation of 1° per image. ALDC-GOL was exposed for 1200 sec per image, 260° of data were collected with a rotation of 1° per image.

2.4.6.2 Diamond Light Source

ALDC-PO4, ALDC-RR and ALDC-SR datasets were collected at beamline IO4 at the Diamond Light Source. ALDC-PO4 and ALDC-RR were exposed for 1 sec per image, 90° of data were collected with a rotation of 0.2° per image. ALDC-SR was exposed for 1 sec per image, 120° of data were collected with a rotation of 0.2° per image.

2.4.7 Data processing

Data were indexed, integrated and scaled using the XDS package [84]. The crystals belonged to the trigonal space group $P3_221$ and contained one molecule per asymmetric unit with a solvent content of 46% by volume.

The CCP4 suite of programmes [85] were used for model building and refinement. Initially a model was built by rigid body refinement using the ALDC structure previously solved by the Fülöp group (4BT2) as a template. Bound compounds were identified from a 2mFo-$\triangle$Fc electron density map [86], new monomer libraries were generated using LIBCHECK through the graphical user interface Sketcher (both part of the CCP4 suite) and built into the map. The structures were refined by alternate cycles of REFMAC [87] and manual refitting using O [88]. Water molecules were added automatically using ARP/wARP [89] at positions of large positive peaks in the difference electron density if they were placed in a suitable hydrogen bonding environment. Restrained isotropic temperature factor refinements were carried out for each individual atom. In the final cycles of refinement, the molecule was subjected to TLS (Torsion, Libration and Skew) refinement. The structure was divided into two major domains by graphical inspection (residues 20-100, 146-165, 224-244 and residues 101-145, 166-223, 245-255), which was confirmed by the TLSMD web server [90].

Chapter 3

Synthesis of ALDC substrate and inhibitors

3.1 Introduction

Mechanistic enzymology studies of the kind presented in this thesis combining structural biology techniques with enzyme kinetic studies require the use of small molecules.

ALDC is an unusual enzyme which operates on both enantiomers of acetolactate to give a single product, (*R*)-acetoin. To gain insight into the molecular interactions that are important in catalysis it is necessary to design and synthesise enzyme inhibitors that resemble intermediates along the reaction pathway. As ALDC has two substrates, inhibitors that separately mimic steps along the (*S*)-acetolactate and (*R*)-acetolactate reaction coordinate are required.

The substrate of ALDC acetolacate is an unstable molecule that can be prepared before use by ester hydrolysis of ethyl 2-acetoxy-2-methylacetoacetate. This acetolactate precursor is

only available commercially as a racemic mixture but as both enantiomers of acetolactate are substrates of ALDC that are turned over at different rates, enzyme kinetic studies require the use of enantiopure acetolactate. There are no commercially available chiral acetolactate precursors, there are three routes in the literature to produce enantiopure acetolactate precursors but they all require multiple steps and are low yielding.

3.2 Aims

3.2.1 Design and synthesis of chiral transition state analogues

The first aim of the work described in this chapter is to synthesise chiral transition state analogues. These compounds will act as inhibitors of ALDC and will be used in future chapters to provide information about the reaction mechanism for both ALDC substrates.

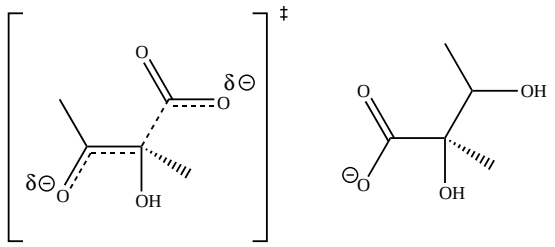
3.2.2 Synthesis of enantiomerically enriched (*S*)-acetolactate

The second aim of the work described in this chapter is to develop a synthetic protocol for the preparation of enantiomerically enriched (*S*)-acetolactate. This enantiomerically pure substrate is not available commercially but it is required for enzyme kinetic studies.

3.3 Synthesis of 2,3-dihydroxy-2-methylbutanoic acids (11)

A two-step synthesis was developed to synthesise the four possible stereoisomers of 2,3-dihydroxy-2-methylbutanoic acid **11**. These compounds were identified as transition state

Figure 3.1: Comparison of the ALDC transition state with synthetic transition state analogues



16 **11**

*The transition state **16** of the ALDC catalysed reaction shown on the left is mimicked by transition state analogues **11** shown on the right. The transition state has a partial negative charge on the ketone that is represented by the negatively charged carboxylate group in the analogue. The transition state also has a partial bond to the leaving carboxylate group, which is mimicked in the analogue by a hydroxyl group.*

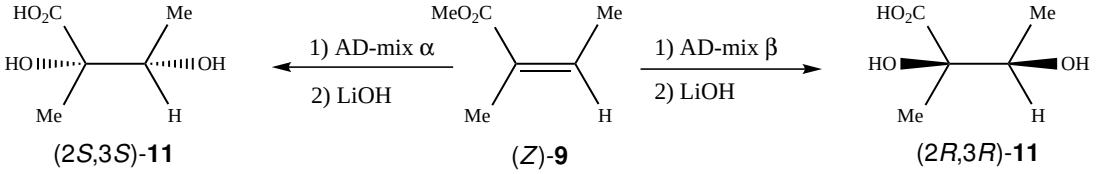
analogues that could be synthesised simply in an enantiomerically enriched form. Figure 3.1 shows the transition state for the acetolactate decarboxylase (ALDC) catalysed reaction and describes how these compounds act as mimics. The transition state has a partial negative charge on the ketone that is mimicked by the negative charge of the carboxylate in the transition state analogue. The leaving carboxylate group in the transition state is mimicked by a hydroxyl group in the analogue.

The synthetic protocol shown in Scheme 3.1 was adapted from Stritzke et al. [65]. The first step is an asymmetric dihydroxylation, Scheme 3.1a shows the (*erythro*)-methyl-2,3-dihydroxy-2-methylbutanoates (2*S*,3*S*)-**10** and (2*R*,3*R*)-**10** were formed from the reaction of methyl angelate (*Z*)-**9** with AD-mix α and AD-mix β respectively. Scheme 3.1b shows the (*threo*)-methyl-2,3-dihydroxy-2-methylbutanoates (2*R*,3*S*)-**10** and (2*S*,3*R*)-**10** which were formed from the reaction of methyl tiglate (*E*)-**9** with AD-mix α and β respectively. The second step is a hydrolysis of the methyl ester to give the desired enantiomerically enriched 2,3-dihydroxy-2-methylbutanoic acids* **11**.

*Preliminary work towards the synthesis of (2*R*,3*S*)-2,3-dihydroxy-2-methylbutanoic acid and (2*S*,3*R*)-2,3-dihydroxy-2-methylbutanoic acid was conducted by the author as part of her MSc dissertation [91].

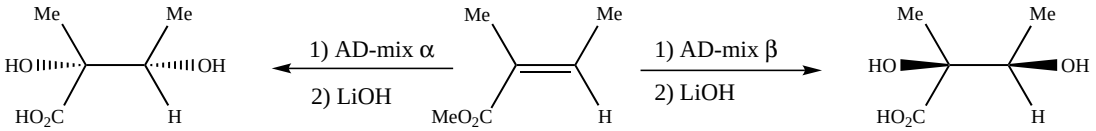
Scheme 3.1: Synthetic route to chiral inhibitors of ALDC

(a) Synthetic route to (*erythro*)-11



(2*S*,3*S*)-**11** (*Z*)-**9** (2*R*,3*R*)-**11**

(b) Synthetic route to (*threo*)-11



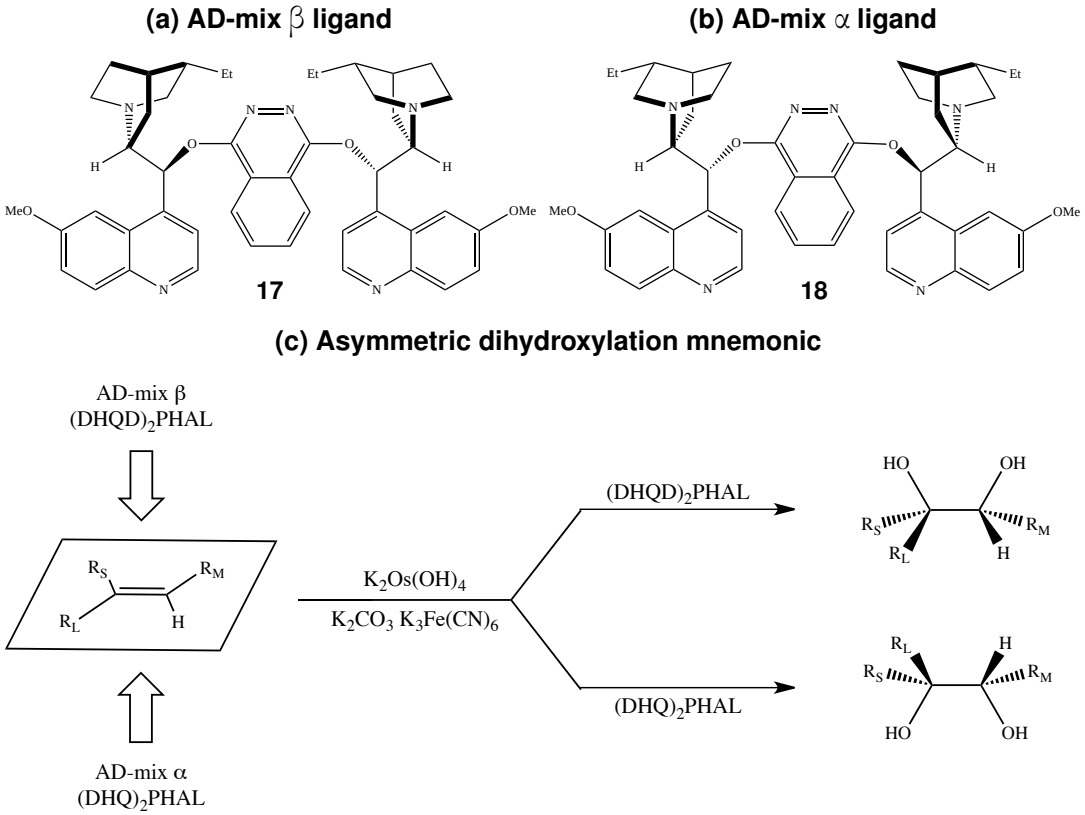
(2*R*,3*S*)-**11** (*E*)-**9** (2*S*,3*R*)-**11**

*Chiral inhibitors of ALDC have been synthesised in a two-step procedure. Firstly, the commercially available alkene (a) methyl angelate (Z)-**9** or (b) methyl tiglate (E)-**9** were asymmetrically dihydroxylated using AD-mix α to give (2S,3S)-**10** and (2R,3S)-**10** or AD-mix β to give (2R,3R)-**10** or (2S,3R)-**10**. This reaction took place in tBuOH/H_2O at 0 °C rising to rt with stirring for 24 h. The second step was an ester hydrolysis, the methyl ester was hydrolysed by stirring with LiOH in H_2O/iPrOH for 8 h at rt.*

3.3.1 2,3-Dihydroxy-2-methylbutanoates (10)

The first step of the protocol generates enantiomerically enriched (*erythro*)-**10** and (*threo*)-**10** products by asymmetrically dihydroxylating methyl angelate (*Z*)-**9** and methyl tiglate (*E*)-**9**, respectively. In this work, the commercially available reagent mixture for the Sharpless asymmetric dihydroxylation (AD-mix) has been used. AD-mix is available in two forms (α and β), which enrich for opposite enantiomers. Both mixtures contain potassium osmate ($K_2OsO_2(OH)_4$), a source of osmium tetroxide (the catalyst of the reaction), potassium ferricyanide ($K_3Fe(CN)_6$), which re-oxidises the catalyst, potassium carbonate (K_2CO_3), which acts as a pH buffer and a chiral ligand to promote selectivity [92]. In AD-mix α the ligand is hydroquinine-1,4-phthalazinediyl ether ($(DHQ)_2PHAL$) and in AD-mix β the ligand is hydroquinidine-1,4-phthalazinediyl ether ($(DHQD)_2PHAL$). The structure of the ligands are shown in Figure 3.2. They are phthalazine adducts of the 'pseudoenantiomeric' cinchona

Figure 3.2: Sharpless asymmetric dihydroxylation



*The sharpless asymmetric dihydroxyation is controlled by chiral ligands, (a) shows $(DHQD)_2PHAL$ **17** which is present in the commercial reagent AD-mix β, (b) shows the pseudoenantiomer $(DHQ)_2PHAL$ **18** found in AD-mix α, (c) shows the mnemonic used to identify the stereochemistry of the product from the reaction.*

alkaloids dihydroquinine (DHQ) and dihydroquinidine (DHQD) and give opposite selectivities [92]. As the ligands are not enantiomers they give slightly different selectivities, with the compounds reported here, the selectivity of AD-mix β is higher.

The stereochemistry of the four compounds has been assigned using the Sharpless asymmetric dihydroxylation mnemonic shown in Figure 3.2c. For methyl angelate (*Z*)-**9** R_L = Me, R_S = CO_2Me, R_M = Me, therefore AD-mix β gives (2*R*,3*R*)-**10** and AD-mix α gives (2*S*,3*S*)-**10**. For methyl tiglate (*E*)-**9** R_L = CO_2Me, R_S = Me, R_M = Me, therefore AD-mix β gives (2*S*,3*R*)-**10** and AD-mix α gives (2*R*,3*S*)-**10**.

Table 3.1: Asymmetric dihydroxylation yields and enantiomeric excess

Compound	Yield	Ee
(2*R*,3*S*)-**10**	63%	75%
(2*S*,3*R*)-**10**	57%	87%
(2*S*,3*S*)-**10**	65%	75%
(2*R*,3*R*)-**10**	51%	85%

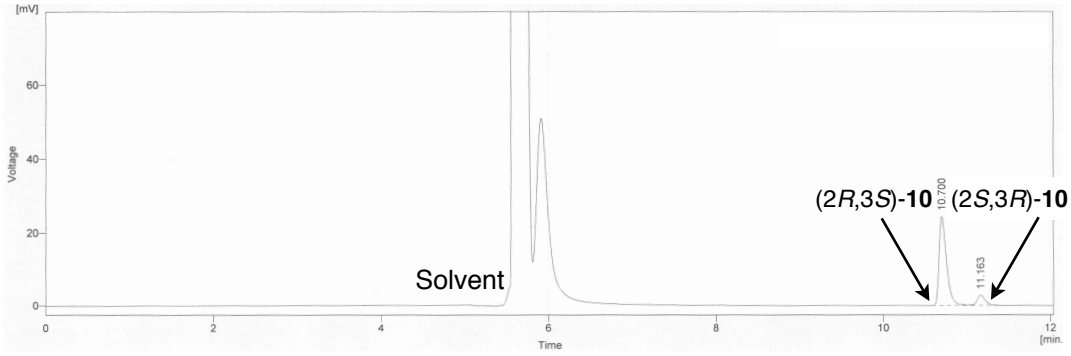
All four compounds have been synthesised in reasonable yields with good ee as summarised in Table 3.1. The ee of the reaction was measured using chiral gas chromatography. Figure 3.3 shows the gas chromatograms for (2*R*,3*S*)-**10** and (2*S*,3*R*)-**10**. Figure 3.3a for (2*R*,3*S*)-**10** shows the major peak is eluted first but this is reversed in Figure 3.3b for (2*S*,3*R*)-**10** where the major peak is eluted second. Figure 3.4 shows the gas chromatograms for (2*S*,3*S*)-**10** and (2*R*,3*R*)-**10**. Figure 3.4a for (2*S*,3*S*)-**10** shows the major peak is eluted second but this is reversed in Figure 3.4b for (2*R*,3*R*)-**10** where the major peak is eluted first.

The data for (*erythro*)-**10** (see Section 2.1.1.1.4) and (*threo*)-**10** (see Section 2.1.1.1.1) are in good agreement with the literature [66–69,71], however there are discrepancies with optical rotations. In this study, the optical rotation of (2*S*,3*S*)-**10** was measured as +27.8° (c=1.1 in $CHCl_3$, ee 75%, 20°C), which is the same sign but slightly higher magnitude than the literature values of +8.7° (c=0.6, $CDCl_3$, 20°C)[71] and +12.3° (c=1.0, $CHCl_3$, ee 99.9%, 20° C) [69]. The optical rotation of (2*R*,3*R*)-**10** was -24.8° (c=1.1, $CHCl_3$, ee 85%, 24°C) and has not been reported in the literature. The assignment of stereochemistry has been confirmed by agreement of the sign of optical rotations with the literature and by crystal structures of both enantiomers of (*erythro*)-**11** (see Section 6.5).

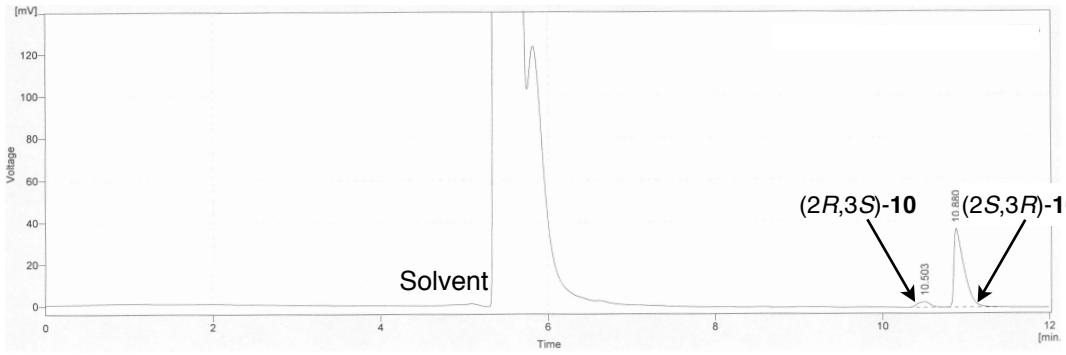
For compound (2*R*,3*S*)-**10**, there are three conflicting optical rotation values reported in the literature -1.0° (c=1.2, $CDCl_3$, 20°C) [67], +1.0° (c=1.0, $CHCl_3$, 20°C) [69] and -

Figure 3.3: Gas chromatograms of (*threo*)-methyl-2,3-dihydroxy-2-methylbutanoate (*threo*)-10

(a) (2*R*,3*S*)-10



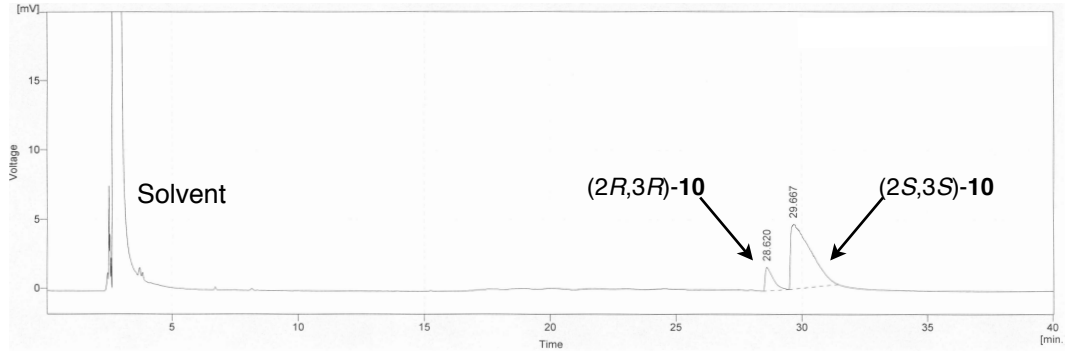
(b) (2*S*,3*R*)-10



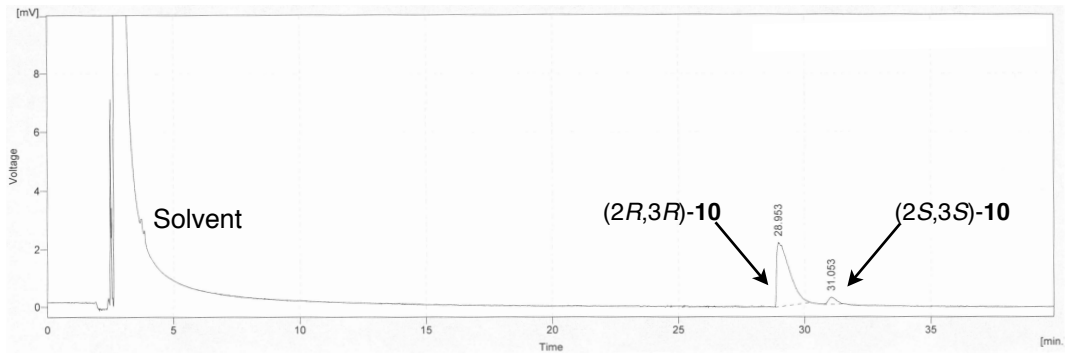
*The enantiomers of (threo)-methyl-2,3-dihydroxy-2-methylbutanoate were separated by chiral gas chromatography, conditions are described in Sections 2.1.1.1.2 and 2.1.1.1.3. Compound (2R,3S)-**10** eluted first followed by (2S,3R)-**10**. (a) The chromatogram for enriched (2R,3S)-**10** 75% ee, the major peak eluted at 10.7 min and the minor peak at 11.2 min. (b) The chromatogram for enriched (2S,3R)-**10** 87% ee, the minor peak eluted at 10.5 min and the major peak at 10.9 min.*

Figure 3.4: Gas chromatograms of (*erythro*)-methyl-2,3-dihydroxy-2-methylbutanoate (*erythro*)-10

(a) (2*S*,3*S*)-10



(b) (2*R*,3*R*)-10



*The enantiomers of methyl-2,3-dihydroxy-2-methylbutanoate were separated by chiral gas chromatography, conditions are described in Sections 2.1.1.1.5 and 2.1.1.1.6. Compound (2R,3R)-**10** eluted first followed by (2S,3S)-**10**. (a) The chromatogram for enriched (2S,3S)-**10** 75% ee, the minor peak eluted at 28.6 min and the major peak at 29.7 min. (b) The chromatogram for enriched (2R,3R)-**10** 85% ee, the major peak eluted at 29.0 min and the minor peak at 31.1 min.*

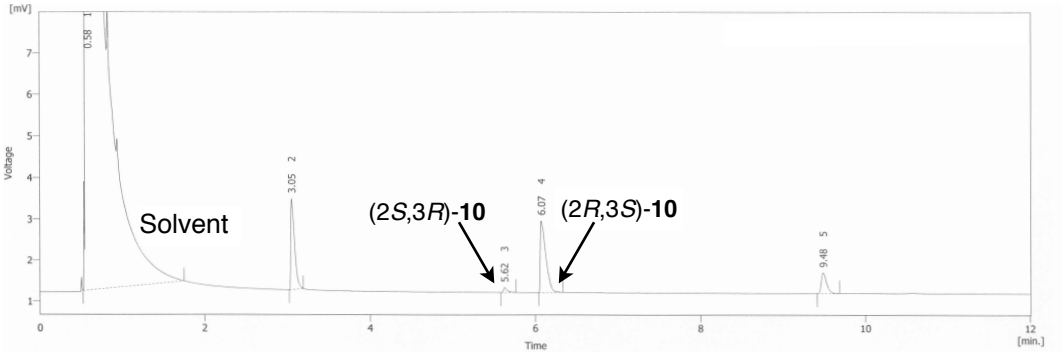
2.5° (neat, 22°C) [70]. Compound (2*R*,3*S*)-**10** made in this study had positive optical rotation values measured on two different occasions of +16.4° (c=1.0, $CHCl_3$, ee 75%, 22° C) and +21.3° (c=1.1, $CHCl_3$, ee 77%, 20°C). For (2*S*,3*R*)-**10**, there is one value reported in the literature +1.0° (c=6.7, $CDCl_3$) [68]. Compound (2*S*,3*R*)-**10** made in this study had two negative optical rotation values measured of -17.9° (c=1.0, $CHCl_3$, ee 87%, 22°C) and -12.9° (c=1.4, $CHCl_3$, ee 84%, 24°C). The values reported here are different from the low literature values. To confirm the stereochemistry was correctly assigned, a small scale kinetic resolution of (2*R*,3*S*)-**10** (ee 75%) with N435 lipase was performed. This kinetic resolution is described in more detail in Section 3.4.1; N435 selectively esterifies the secondary alcohol with *R* configuration, which in this case is (2*S*,3*R*)-**10**. Therefore, if the stereochemistry had been correctly assigned the ee would increase and if it had been incorrectly assigned the ee would decrease. Figure 3.5 shows the gas chromatogram that was obtained from that reaction and an ee of 92% was measured. This increase in ee demonstrates that the minor isomer is being esterified confirming the assignment of stereochemistry.

3.3.2 Improving the yield and purity

This reaction has been optimised to improve the yield and purity of the recovered compounds. Compound **10** is rather polar so it was difficult to recover from the aqueous phase. The yield was improved by extracting from the aqueous layer more times and bypassing the addition of water before the work up.

The initial reaction protocol included the addition of methanesulfonamide but this reagent persisted as an impurity even after column purification. It was identified by NMR, giving rise to a singlet at 3.08 ppm in proton NMR and a peak at 43.1 ppm in carbon NMR. Junttila et al. [93] showed that methanesulfonamide acceleration can occur via two mechanisms: a

Figure 3.5: Gas chromatogram of the kinetic resolution of (2*R*,3*S*)-methyl-2,3-dihydroxy-2-methylbutanoate (2*R*,3*S*)-10



*This gas chromatogram was produced from a sample of a small scale kinetic resolution reaction of enantiomerically enriched (2R,3S)-**10**. A sample of the reaction was diluted in EtOAc and filtered through cotton wool and injected onto a CP-ChiraSil-DEX CB 25 m × 0.25 mm × 0.25 μm, with hydrogen as the carrier gas, the initial oven temperature was set to 90 °C and the temperature was ramped at 4 °C/min to 115 °C and held for 30 min. The reaction starting material had been assigned as (2R,3S)-**10** and had an ee of 75%. This gas chromatogram shows that the ee has increased to 92%.*

general acid-catalysed hydrolysis and as a cosolvent of hydroxide ions. The acid-catalysed effect is significant when it comes to aromatic compounds and acting as a cosolvent is important for intermediate osmate esters that are non-polar. In this case the compound is not aromatic or non-polar so the necessity of methanesulfonamide addition was investigated. The absence of methanesulfonamide had no adverse effect on the reaction and it led to a cleaner product.

3.3.3 2,3-Dihydroxy-2-methylbutanoic acids (11)

The second stage of the reaction was an ester hydrolysis (Scheme 3.1). The procedure in Stritzke et al. [65] was adapted: the ester was stirred in LiOH rather than KOH and the solvent was a 1:1 mixture of IPA:H_2O rather than EtOH:H_2O. The hydrolysis conditions were milder, stirring at rt overnight rather than at 50 °C for 36 h. It was found that the reaction

Table 3.2: Ester hydrolysis yields

Compound	Yield
(2*R*,3*S*)-**11**	67%
(2*S*,3*R*)-**11**	51%
(2*S*,3*S*)-**11**	72%
(2*R*,3*R*)-**11**	61%

went to completion under the milder conditions. The work up conditions were also modified: in the published protocol the reaction was worked up by addition of acid and saturated with NaCl before an extraction with Et_2O. In this case it was found that removal of the NaCl step and more extractions resulted in an improved yield. The yields obtained were reasonable and are summarised in Table 3.2.

These four compounds have been previously reported but not fully characterised; the NMR data for both *erythro* and *threo* agree with the literature [73,75]. The optical rotation of (2*R*,3*S*)-**11** was measured to be -0.5° (c=1.05, MeOH, ee 75%, 29°C), which has a different sign than the literature value of +5.7° (c=4.6, H_2O, 27°C) [74] but it is not comparable as it is measured in a different solvent. Attempts to measure the optical rotation in the same solvent were unsuccessful, as when the compound is dissolved in water the solution is more viscous and could not be placed in the polarimeter cell without generating bubbles. The optical rotation of (2*S*,3*R*)-**11** was measured as +0.7° (c=1.0, MeOH, ee 84%, 25°C) and has not been previously reported in the literature. The optical rotation of compound (2*R*,3*R*)-**11** was measured as -4.7° (c=1.0, MeOH, ee 85%, 25°C) and has not been previously reported. The optical rotation for compound (2*S*,3*S*)-**11** was measured as +3.5° (c=1.2, MeOH, ee 75%, 24°C). There is a literature value reported for this compound of -2.25° (c=3.67, MeOH) [75]. The value measured in this study does not agree with the reported value; in the paper referenced, Kaneko et al. [73] derived this compound from a natural product and confirmed it to be an enantiomer of (*erythro*)-2,3-dihydroxy-2-methylbutanoic acid (*erythro*)-**11**. The

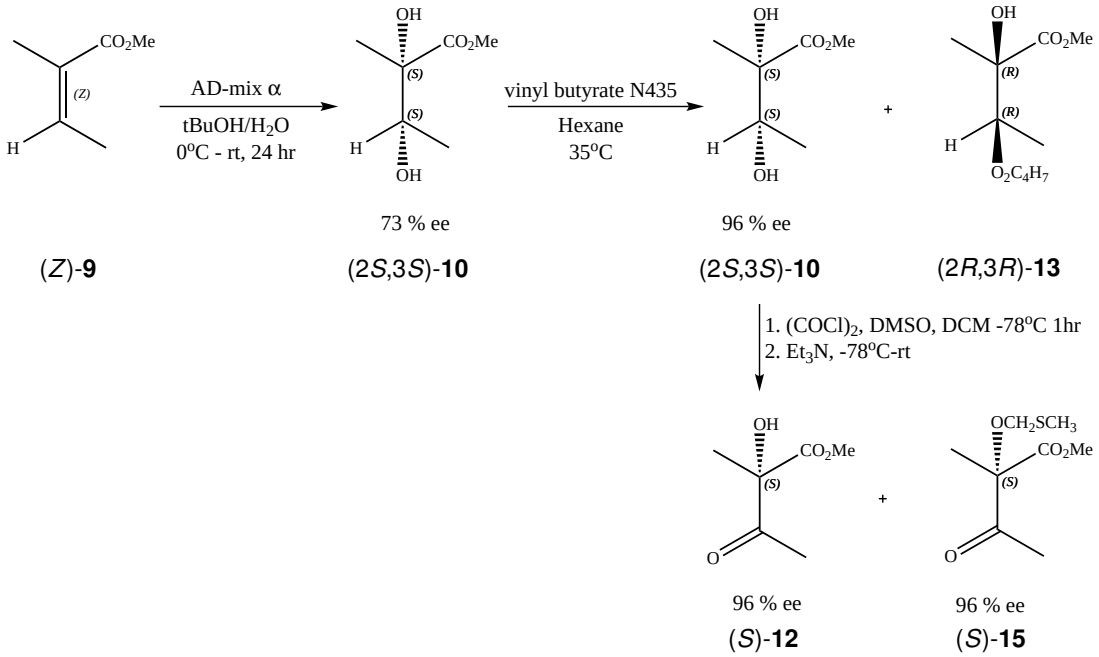
stereochemistry was putatively assigned as (2*S*,3*S*)-**11** based on chiral HPLC, however this is inconsistent with the optical rotation measured in this study of (2*S*,3*S*)-**11**. As the stereochemistry assignment in this study has been confirmed by crystal structures (Section 6.5) it appears that Kaneko et al. [73] incorrectly assigned stereochemistry.

3.4 Synthesis of enantiomerically enriched (*S*)-acetolactate (*S*)-3

Optically pure (*S*)-acetolactate (*S*)-**3** was required for use in ALDC activity studies. Acetolactate **3** is an unstable β-keto acid that has to be prepared immediately before use. Ethyl 2-acetoxy-2-methylacetoacetate **19** is an acetolactate **3** precursor that is available commercially but only as a racemate. As both enantiomers of acetolactate **3** are substrates of ALDC it was necessary to produce optically pure (*S*)-acetolactate (*S*)-**3** for use in enzyme assays to reliably determine kinetic parameters.

A new synthetic route to enantiomerically enriched (*S*)-acetolactate (*S*)-**3** has been developed. There are three routes to an enantiomerically enriched acetolactate ester in the literature [55,76,79]. Two of the routes [55,79] involve 5 steps including a low yielding resolution with quinine. The alternative route [76] takes 6 steps and involves the formation of diastereomeric dioxolanones from lactic acid that can be separated on a column and converted into both optically pure enantiomers of ethyl acetolactate. The route presented here (Scheme 3.2) is superior to the previous routes, it is shorter taking only 3 steps and can be used to make both enantiomers of methyl acetolactate **12**; only the route to (*S*)-methyl acetolactate (*S*)-**12** is reported. The first stage involves the same asymmetric dihydroxylation of methyl angelate with AD mix α to give enantiomerically enriched (2*S*,3*S*)-methyl-2,3-dihydroxy-2-

Scheme 3.2: Synthetic route to (*S*)-methyl acetolactate (*S*)-12



*Enantiomerically enriched (S)-methyl acetolactate has been synthesised in a three-step procedure. Firstly, the commercially available alkene methyl angelate (Z)-**9** was asymmetrically dihydroxylated with AD-mix α to give 73% ee (2S,3S)-**10**. Secondly, the ee of (2S,3S)-**10** was increased by a kinetic resolution: N435 lipase esterifies the minor enantiomer (2R,3R)-**10** at the 3 position to give (2R,3R)-**13**. Compound (2S,3S)-**10** was separated from the mixture by flash column chromatography and used as the substrate for a Swern oxidation, which oxidised the secondary alcohol to a ketone to give (S)-**12**, which could be purified by column chromatography to separate the desired product from the methylthiomethylation side product (S)-**15**.*

methylbutanoate (2*S*,3*S*)-**10**, discussed in Section 3.3.1. To make optically pure methyl acetolactate **12**, the second reaction is a kinetic resolution [69] and the final reaction is a Swern oxidation [76].

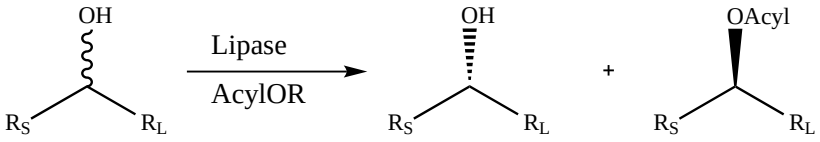
3.4.1 Kinetic resolution of (2*S*,3*S*)-methyl-2,3-dihydroxy-2-methylbutanoate (2*S*,3*S*)-10

Moen et al. [69] reported the resolution of (2*S*,3*S*)-methyl-2,3-dihydroxy-2-methylbutanoate (2*S*,3*S*)-**10** from a racemic mixture using the commercial lipase N435 (Novozyme). N435 selectively esterifies secondary alcohol centres according to Kazlauskas rule [94] (shown in Figure 3.6). When applied in this case, the secondary alcohol with *R* configuration is esterified. Compound (2*R*,3*R*)-**10** is esterified and (2*S*,3*S*)-**10** is left unchanged. This reaction has been used in this study to increase the ee of enriched (2*S*,3*S*)-**10** from 75% to >95%. The enrichment reaction was monitored by GC as shown in Figure 3.7a, which allowed the ee to be calculated during the reaction. When the desired ee was reached the reaction could be worked up and the desired compound (2*S*,3*S*)-**10** purified by column chromatography. Figure 3.7b shows the gas chromatogram obtained after the compound had been purified, the ee of the purified sample was 96%.

3.4.2 Oxidation of the secondary alcohol

The final reaction is an oxidation of the secondary alcohol to give (*S*)-methyl acetolactate (*S*)-**12**. Two different methods were attempted: a Swern oxidation, which has been reported for ethyl acetolactate [76] and a Dess-Martin oxidation.

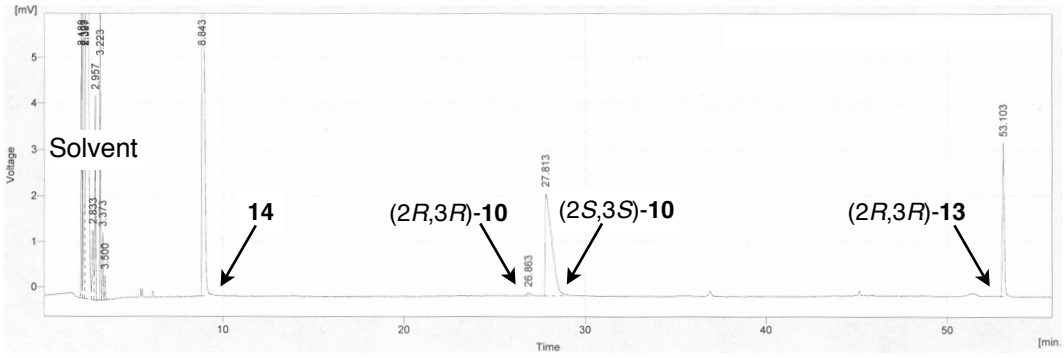
Figure 3.6: Kazlauskas rule for selectively of lipases for secondary alcohols



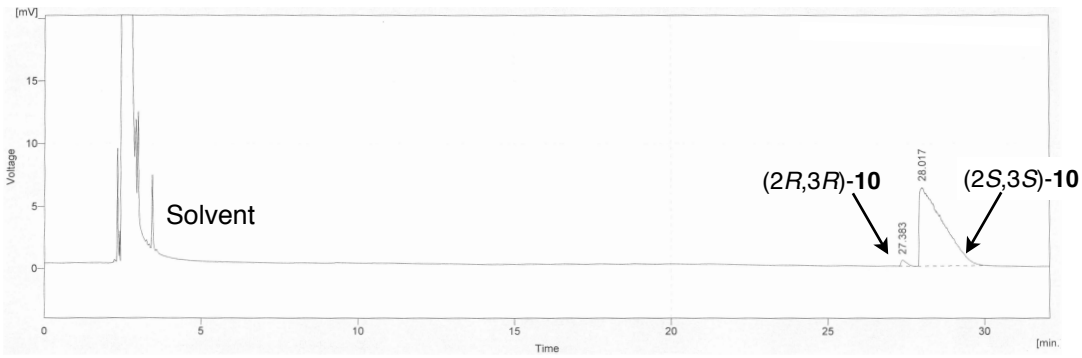
Kazlauskas rule describes the selectivity of lipases at esterifying secondary alcohols. When there are two groups, one that is smaller (R_S) and one that is larger (R_L), the lipase can differentiate between the two enantiomers and kinetically resolve them.

Figure 3.7: Gas chromatograms of the kinetic resolution of (2*S*,3*S*)-10

(a) Gas chromatogram of the reaction mixture



(b) Gas chromatogram of the purified product



*The kinetic resolution of (2S,3S)-**10** was monitored by chiral GC. (a) Chromatogram that was obtained during a reaction, the peak observed at 8.8 min is that of butyric acid **14** a side product formed from the degradation of vinyl butyrate. At 26.9 and 27.8 min (2R,3R)-**10** and (2S,3S)-**10** were eluted and at 53.1 min (2R,3R)-**13** was eluted. (b) The chromatogram of (2S,3S)-**10** after purification, (2R,3R)-**10** was eluted at 27.4 min followed by (2S,3S)-**10** at 28.0 min. The ee of this sample was 96%.*

3.4.2.1 Swern oxidation

The Swern oxidation oxidises the secondary alcohol at the 3 position to give the ketone in the product. This oxidation has been reported for the ethyl ester of 2-methyl-2,3-dihydroxybutanoic acid in a 47% yield [76]. It was reported that under stoichiometric conditions only partial conversion was obtained but when adding excess reagents, methylthiomethylation of the tertiary alcohol was observed. (*S*)-Methyl acetolactate (*S*)-**12** was recovered in 69% yield with minimal presence of the side product **15** (not recovered) on a 3 g scale. When the reaction was carried out on lower scales, the yield was lower and the presence of side product increased. The desired compound was challenging to purify due to a low boiling point of 74-76 °C [78]. Purification was achieved by column purification on silica gel using a solvent system of low boiling point solvents pentane and diethyl ether. Solvent was removed carefully using a rotary evaporator under low vacuum.

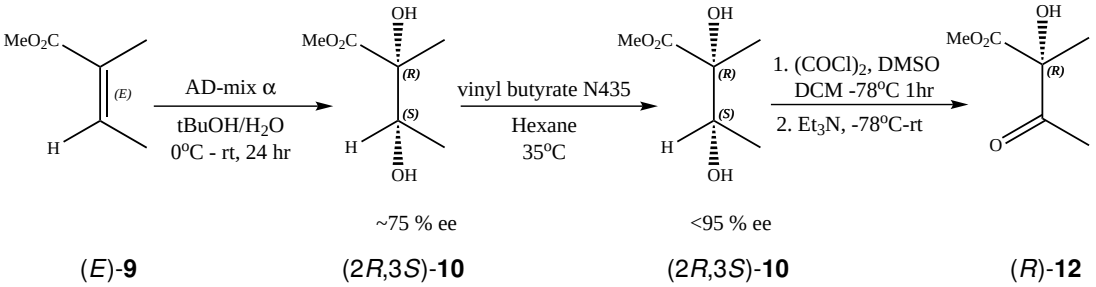
3.4.2.2 Dess-Martin oxidation

An alternative oxidation method was tried. The Dess-Martin oxidation was performed on a small scale. On the first occasion crude NMR showed that the product had formed but the material contained impurities and the reaction had not gone to completion. Attempts to reproduce the reaction were unsuccessful and it was decided to continue with the established method of Swern oxidation.

3.4.3 (*R*)-Methyl acetolactate

The same three-step synthetic protocol described here could be used to generate (*R*)-methyl acetolactate (*R*)-**12** using methyl tiglate and AD mix α to give enriched (2*R*,3*S*)-methyl-

Scheme 3.3: Synthetic route to (*R*)-methyl acetolactate (*R*)-12



~75 % ee <95 % ee

(*E*)-**9** (2*R*,3*S*)-**10** (2*R*,3*S*)-**10** (*R*)-**12**

*Enantiomerically enriched (R)-methyl acetolactate could be synthesised via the same three-step procedure used to synthesise (S)-acetolactate (S)-**12**. Firstly, the commercially available alkene methyl tiglate (E)-**9** could be asymmetrically dihydroxylated with AD-mix α to give 75% ee (2R,3S)-**10**. Secondly, the ee of (2R,3S)-**10** could be increased by a kinetic resolution and used as the substrate for a Swern oxidation to give (R)-**12**.*

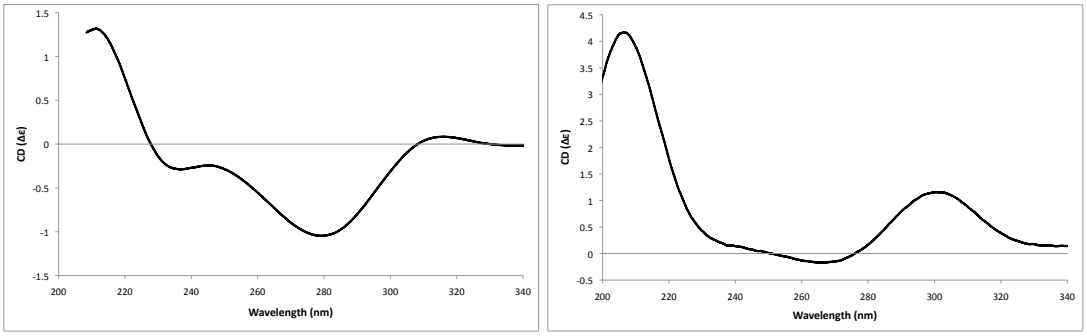
2,3-dihydroxy-2-methylbutanoate (2*R*,3*S*)-**10**, which when oxidised would give (*R*)-methyl acetolactate. This is illustrated in Scheme 3.3.

3.4.4 Hydrolysis of (*S*)-methyl acetolactate (*S*)-12

It is reported in the literature [41] that acetolactate **3** is unstable under acidic conditions so ester hydrolysis is conducted under alkaline conditions. It is also described in the literature that acetolactate **3** will racemise under alkaline conditions [56]. The mechanism of this racemisation has been determined to occur by a tertiary ketol rearrangement [54]. To ensure that any (*S*)-acetolactate (*S*)-**3** that was formed by hydrolysis was still enantiomerically enriched the hydrolysis reaction was monitored using circular dichroism. Figure 3.8 shows the CD spectra of (*S*)-methyl acetolactate (*S*)-**12** and (*S*)-acetolactate (*S*)-**3**. The CD spectrum of both (*S*)-**3** [95] and (*S*)-**12** [79] have been reported in the literature and the results presented here agree, (*S*)-**3** has a positive peak at 300 nm. At this wavelength a molar ellipticity value of 3350 deg cm^2 $dmol^{-1}$ was measured for a 96% ee sample, agreeing with the reported value of 3600 deg cm^2 $dmol^{-1}$. The CD spectrum for (*S*)-**12** was performed

Figure 3.8: CD spectra of (*S*)-methyl acetolactate (*S*)-12 and (*S*)-acetolactate (*S*)-3

(a) (*S*)-Methyl acetolactate (*S*)-12 **(b) (*S*)-Acetolactate (*S*)-3**



*The CD spectra of enantiomerically enriched (S)-methyl acetolactate (S)-**12** and (S)-acetolactate (S)-**3** were recorded on a J-715 spectropolarimeter. (a) The spectrum of methyl acetolactate has a positive peak at 211 nm, two negative peaks at 237 nm, 279 nm and a further positive peak at 316 nm. (b) The spectrum of acetolactate has a positive peak at 207 nm, a negative peak at 267 nmm and a further positive peak at 300 nm.*

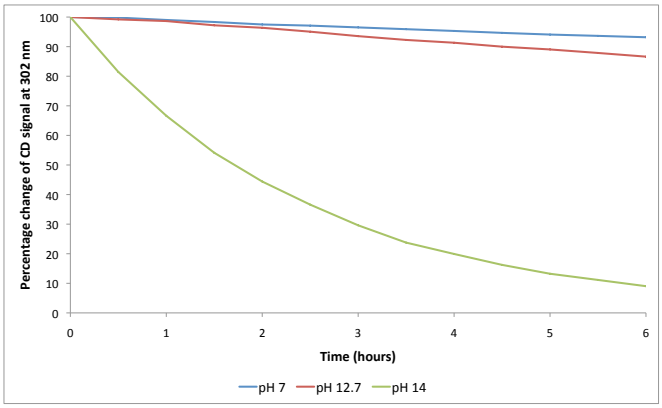
in a different solvent so the magnitude of values are different but the four peaks reported are observed at the same wavelengths. It was observed that the hydrolysis of (*S*)-**12** in the presence of an excess of NaOH had proceeded to completion in the time it took to invert the curvette 3 times, place into the CD spectropolarimeter and complete a scan. The CD signal was monitored in alkaline conditions for 10 minutes and no racemisation had occured. It was concluded that a 5 min stir followed by dropping the pH would allow the hydrolysis to go to completion and the product formed would not undergo racemisation.

3.4.4.1 Stability at different pH

A more thorough investigation into the stability of (*S*)-acetolactate (*S*)-**3** under 3 different pH conditions was conducted. Wavelength scans were conducted every 10 or 30 min at pH 7, 12.7 and 14. Figure 3.9 shows the degradation of (*S*)-**3** over time. As expected (*S*)-**3** degraded most rapidly at pH 14 due to racemisation, at pH 12.7 the rate of degradation was slower and most likely due to a mixture of racemisation and decarboxylation. At pH 7 the

degradation was due to decarboxylation and (*S*)-**3** at rt degraded by 6.8% in 6 hours.

Figure 3.9: Stability of (*S*)-acetolactate (*S*)-3 at different pH



*The stability of (S)-acetolactate (S)-**3** was studied at pH 14 (green), pH 12.7 (red) and pH 7 (blue). This figure shows the change in CD_{302nm} over 6 hours, at pH 14 the CD signal had decreased by 91.0%, at pH 12.7 the CD signal had decreased by 13.4% and at pH 7 the CD signal had decreased by 6.8%.*

3.5 Conclusion

The work detailed in this chapter has been fundamental to future chapters. A series of chiral transition state analogues have been synthesised in reasonable yields with good ee. The compounds have been fully characterised and are in good agreement with the literature.

A new synthetic route to the natural substrate of ALDC, (*S*)-acetolactate (*S*)-**3** has been developed which is shorter than the existing routes described in the literature. Although (*R*)-acetolactate (*R*)-**3** has not been synthesised during this work, the protocol described here could be used to produce it.

Circular dichroism has been successfully used to monitor the conversion of (*S*)-methyl acetolactate (*S*)-**12** into (*S*)-acetolactate (*S*)-**3** and also to study its stability over time.

Chapter 4

Cloning, expression and purification of *B. subtilis* AlsD and BdhA

4.1 Introduction

Bacillus subtilis is a well studied gram positive bacterium that is naturally found in soil and vegetation and is capable of fermenting pyruvate to 2,3-butanediol. The genome of *B. subtilis* is sequenced and well characterised, including the genes involved in the butanediol fermentation pathway: *alsS* encoding acetolactate synthase, *alsD* encoding ALDC and *bdhA* encoding acetoin reductase. The first two genes *alsS* and *alsD* are contained in the *alsSD* operon which is under the regulation of a LysR type transcriptional regulator encoded by the *alsR* gene [96]. The final gene, *bdhA* was originally misannotated as a protein similar to L-iditol 2-dehydrogenase and named *ydjL*. However it has since been identified as an acetoin reductase due to amino acid sequence similarity and was confirmed to be the primary acetoin reductase in *B. subtilis*. The gene has since been renamed to *bdhA* to reflect its function as an acetoin reductase or butanediol dehydrogenase [97]. The *bdhA* gene is encoded sep-

arately from the *alsSD* operon but it has been shown to be under the indirect control of AlsR [98].

4.2 Aims

4.2.1 Cloning, expression, purification and site directed mutagenesis of AlsD

The first aim of the work described in this chapter is to clone *alsD* into an expression vector and generate single point mutations of active site residues. Conditions for over-expression and purification of recombinant native and mutant AlsD will be optimised. Purified recombinant AlsD will be used in kinetic studies, described in Chapter 5, to give an insight into the catalytic activity of ALDC.

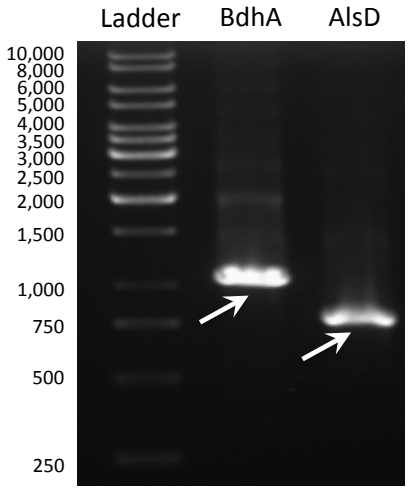
4.2.2 Cloning, expression and purification of BdhA

The second aim of the work described in this chapter is to clone *bdhA* into an expression vector, develop conditions for the over-expression of BdhA in *E. coli* and establish purification procedures. BdhA is required for the development of a coupled assay for ALDC activity which is described in Chapter 5.

4.3 Molecular cloning of *B. subtilis bdhA* and *alsD*

Two *Bacillus subtilis* genes, *alsD* and *bdhA* (encoding acetolactate decarboxylase and acetoin reductase, respectively), were cloned into seperate *E. coli* expression vectors for further

Figure 4.1: Amplification of *B. subtilis bdhA* and *alsD* by PCR



Two B. subtilis genes (bdhA and alsD) were amplified from genomic DNA. The PCR products were run on a 0.8% (w/v) agarose gel and are labelled with arrows, bdhA is 1041 bp and alsD is 832 bp.

studies. BdhA was required for the development of a coupled assay (see Section 5.4) and AlsD was required for mutagenesis studies.

4.3.1 Generation of expression vectors containing *bdhA* and *alsD*

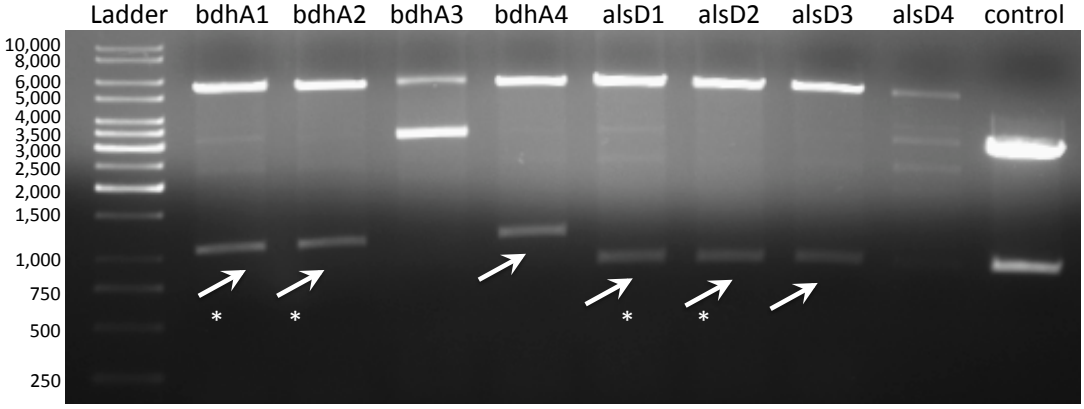
Bacillus subtilis bdhA and *alsD* were amplified from genomic DNA as described in Section 2.2.2.2. The PCR primers (Table 2.2) were designed to introduce a BamHI and XhoI restriction site at the 5' and 3' end of the gene sequence, respectively. The PCR products were analysed on a 0.8% (w/v) agarose gel (Figure 4.1) which showed both products were successfully amplified. The PCR products were double digested with BamHI and XhoI and ligated into an expression vector (pET28a) as described in Section 2.2.2.7. *E. coli* Top10 were transformed with ligation mix and grown on agar plates containing kanamycin to select for bacteria that had taken up the plasmid. The ligations were successful with >200 colonies per plate. Four colonies containing pET28bdhA and four colonies containing pET28alsD

were grown in 5 ml of LB overnight and the plasmids were isolated by mini-prep to give 8 total plasmid preparations, four of the construct pET28bdhA and four of pET28alsD.

4.3.2 Verification of constructs

The constructs described in Section 4.3.1 were confirmed by diagnostic restriction digest and DNA sequencing. Figure 4.2 shows the results of a diagnostic digest, the four samples of pET28alsD (lane 2-5), four samples of pET28bdhA (lane 6-9) and a positive control were digested with BamHI and XhoI. The arrows on the gel indicate that 6 of the 8 plasmids contained a fragment of the correct molecular weight (1061 bp for bdhA and 832 bp for alsD). Four plasmids were sent for sequencing (denoted by asterisks) and they were found to contain the correct gene, in frame and with no mutations. All future work was conducted on the verified samples labelled alsD1 and bdhA1.

Figure 4.2: Diagnostic restriction digest of pET28bdhA and pET28alsD



Expression vectors were subjected to a diagnostic restriction digest, three of the four samples of bdhA that were digested contained an insert of the correct size and three of the four samples of alsD contained an insert of the correct size (labelled with arrows). The samples marked with an asterisk (∗) were confirmed by sequencing to contain the desired inserts in frame with no mutations.

4.3.3 Site-directed mutagenesis

Active site mutants of *alsD* were generated for future studies to identify the residues involved in catalysis. As there is no crystal structure of AlsD, the residues were identified from the crystal structure of *Bacillus brevis* ALDC. Four residues were identified as targets for mutation: Thr58, Glu65, Arg145 and Glu253. The *B. brevis* sequence was aligned against the *B. subtilis* sequence using the alignment tool of UniProt [99] and corresponding residues were identified. Table 4.1 summarises the residue numbering and the mutations that were made. Each residue was replaced with an alanine, and also with an amino acid sharing more similar physicochemical properties. Glutamate was mutated to glutamine, that has a similar shape and hydrogen bonding properties but different charge. Threonine was mutated to serine, which has similar hydrogen bonding properties but is smaller. Arginine was mutated to lysine, another positively charged amino acid.

The 8 single point mutations were introduced into the *alsD* gene in the pET28alsD1 construct using the Agilent QuikChange Lightning kit. Primers were designed using the Agilent online primer design tool, the primers were checked, altered if necessary and ordered from IDT. All 8 mutants were made successfully and their identity confirmed by DNA sequencing*.

Table 4.1: Mutations of *B. subtilis* and the corresponding *B. brevis* residues

***B. brevis* residue**	***B. subtilis* residue**	**Mutations made**
T58	T55	T55A, T55S
E65	E62	E62A, E62Q
R145	R142	R142A, R142K
E253	E251	E251A, E215Q

*E62A, E62Q, R142A, R142K, T55S and E251A were prepared by Claire Dow during her MSc project [100].

4.4 Expression

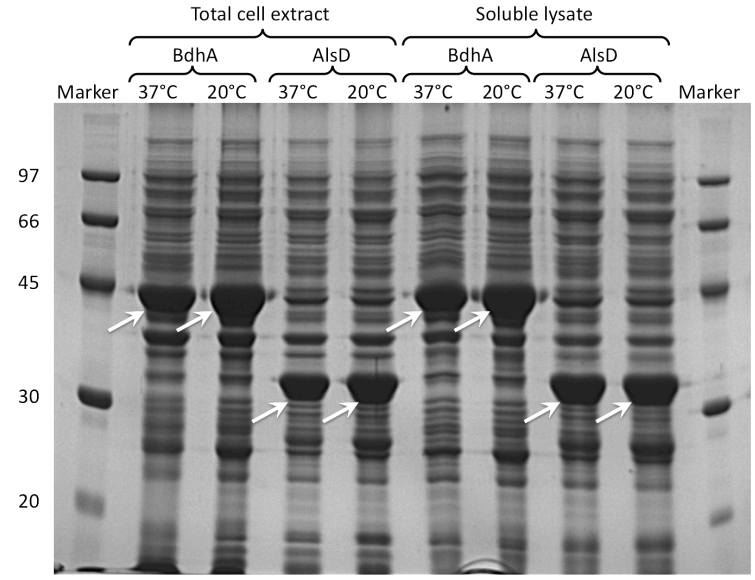
The expression vector used in this study (pET28a) was selected because it contains a suitable multiple cloning site to insert the genes, a kanamycin resistance gene allowing selection of positive clones, T7 promoter allowing expression to be selectively induced by the addition of IPTG and an N-terminal hexa-histidine tag for purification.

The T7 expression system was originally developed by Studier et al. [101] for the over-expression of recombinant proteins in *E. coli*. The expression of the recombinant protein is under the control of the T7 promoter, and can be induced by the addition of IPTG, when the cells have grown to mid-log phase. T7 RNA polymerase is not naturally present in *E. coli*, but the expression strain used in this study is a DE3 lysogen, which encodes T7 polymerase gene under the control of a modified *lac* operator. In the absence of IPTG, T7 bacteriophage RNA polymerase is repressed. When IPTG (an analogue of lactose) is added the expression of T7 RNA polymerase is activated, which indirectly activates the expression of the recombinant proteins encoded on the plasmid.

4.4.1 Expression temperature trials

Constructs pET28bdhA1 and pET28alsD1 were used to transform chemically competent *E. coli* B834 (DE3) pRare cells, and BdhA and AlsD were over-expressed at two different temperatures. Cells were grown at 37 °C until they reached mid-log phase (OD_{600} of 0.6-0.8), induced with 0.5 mM IPTG and incubated for either 3 h at 37 °C or overnight at 20 °C. Figure 4.3 shows that both proteins had a high level of over-expression in the soluble cellular fraction under both conditions. As both conditions were successful the quicker and more convenient condition of IPTG induction at 37 °C was chosen for future expression experiments.

Figure 4.3: Over-expression of *B. subtilis* BdhA and AlsD in *E. coli* at 20°C and 37°C



The expression of BdhA and AlsD were trialled at two different temperatures, 37°C for 3 h and 20°C overnight. The total cell extract and soluble lysate for both proteins at both temperatures are shown. The over-expressed proteins are marked with arrows.

4.5 Protein purification

It was necessary to purify the recombinant proteins from the mixture containing *E. coli* proteins before further use. As both proteins contain a hexa-histidine tag (his-tag) the first technique employed was immobilised metal affinity chromatography. The his-tag has a high affinity for divalent metal ions so proteins containing a his-tag will bind to a nickel column whereas non-tagged proteins will not. The tagged proteins can be eluted from the column by washing with a buffer containing high concentrations of imidazole.

4.5.1 Purification of BdhA

BdhA was purified in a single step to give a purity that was suitable for further studies.

4.5.1.1 Immobilised metal affinity chromatography

Cells were harvested from 1 L cultures (grown as described above) by centrifugation at 6,000 rpm, cell pellets were re-suspended in 30 ml buffer A (20 mM HEPES, 0.5 M NaCl, 10% glycerol, 10 mM imidazole), lysed by sonication at 70% power for 3 $\times$ 30 sec, and extracts were clarified by centrifugation at 20,000 rpm. The soluble lysate from the BdhA expression was loaded onto a 5 ml chelating sepharose column (GE Healthcare) charged with nickel. The majority of *E. coli* proteins have no/low affinity for divalent metal ions and flow through the column. The column was then washed with 10% buffer B (containing 50 mM imidazole) then 20% buffer B (containing 100 mM imidazole) to remove any weakly bound *E. coli* proteins. Pure his-tagged protein was eluted with 100 % buffer B (20 mM HEPES pH 8.0, 0.5 M NaCl, 500 mM imidazole). Figure 4.4 shows the chromatogram and SDS-PAGE gel of the purification. The eluted protein was judged pure enough for use in the development of a coupled assay.

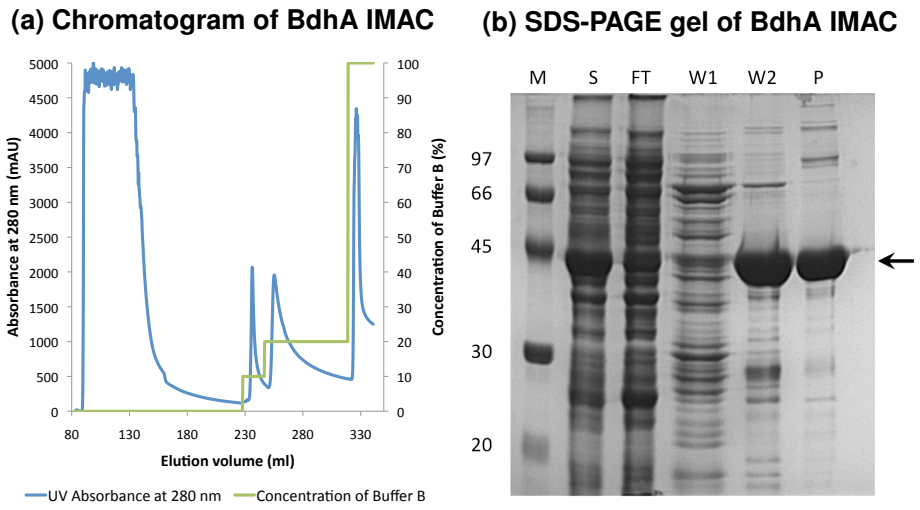
4.5.2 Purification of AlsD

The purification of AlsD was challenging; the following sections describe the purification techniques that were utilised in attempts to purify AlsD and the identification of contaminating polypeptides.

4.5.2.1 Immobilised metal affinity chromatography

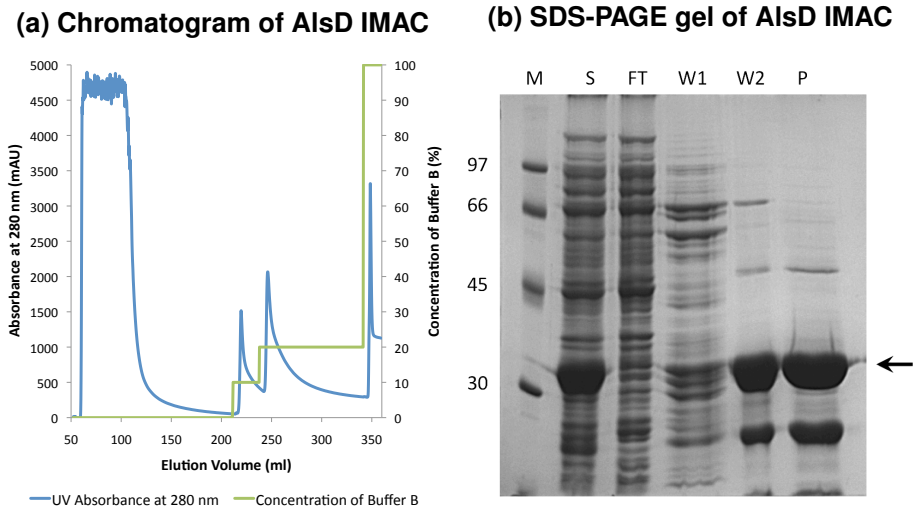
AlsD was purified from the total soluble lysate by IMAC under the same conditions described for BdhA (Section 4.5.1.1). The chromatogram and SDS-PAGE gel of the purification are shown in Figure 4.5. In this case, the protein that was eluted contained other contaminants

Figure 4.4: Purification of *B. subtilis* BdhA by IMAC



B. subtilis BdhA was purified by immobilised metal affinity chromatography (a) shows the chromatogram and (b) shows an SDS-PAGE gel of the purification, M (molecular weight markers), S (soluble lysate), FT (flow through), W1 (10% buffer B wash), W2 (20% buffer B wash), P (eluted protein), the position of BdhA on the gel is marked by an arrow.

Figure 4.5: Purification of *B. subtilis* AlsD by IMAC



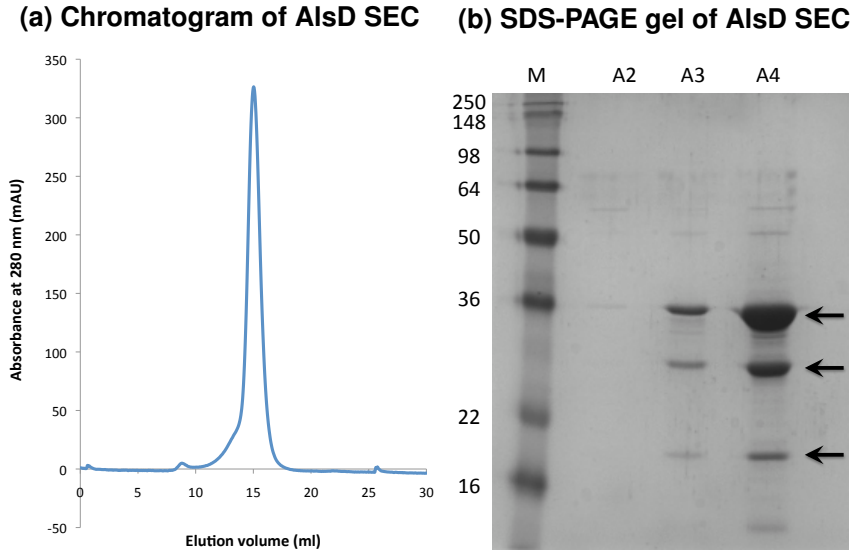
B. subtilis AlsD was purified by immobilised metal affinity chromatography (a) shows the chromatogram and (b) shows an SDS-PAGE gel of the purification, M (molecular weight markers), S (soluble lysate), FT (flow through), W1 (10% buffer B wash), W2 (20% buffer B wash), P (eluted protein), the position of AlsD on the gel is marked by an arrow.

and required further purification.

4.5.2.2 Gel filtration chromatography

As the contaminants were smaller than AlsD on SDS PAGE (ignoring potential oligomers) it was decided to attempt further purification using gel filtration chromatography. Gel filtration chromatography is a technique that purifies proteins based on size. The column is made of porous beads, larger molecules are unable to diffuse into the pores of the beads and are eluted first, whereas smaller molecules can diffuse into the pores and are eluted later. Figure 4.6 shows the chromatogram and SDS-PAGE gel obtained from a gel filtration column. The column was unable to resolve the mixture and the protein that was eluted was no purer after the gel filtration column.

Figure 4.6: Purification of *B. subtilis* AlsD by SEC



Size exclusion chromatography of AlsD, (a) shows the chromatogram (b) shows an SDS-PAGE gel of the eluted protein. M (molecular weight markers), A2 (small peak), A3 (shoulder), A4 (main peak). The purification was unsuccessful, the shoulder and main peak contained the same three products (marked by arrows) seen after IMAC.

4.5.2.3 Ion exchange chromatography

The third purification technique used was ion exchange chromatography. Ion exchange chromatography is an affinity chromatography technique to separate proteins based upon ionic interactions. It was necessary to calculate the theoretical isoelectric point (PI) of AlsD to determine whether to use anion or cation exchange chromatography, as the pH needs to be at least one unit above or below the PI respectively. The PI of AlsD is 4.79 as calculated using ProtParam [102] from the ExPASy Bioinformatics Research Portal, therefore anion exchange chromatography was most suitable. In anion exchange chromatography the protein has an overall negative charge, and binds to a resin containing positively charged groups. A gradient of increasing salt concentration is applied to the column to the elute the bound proteins. In this case, anion exchange chromatography was unable to further purify AlsD.

4.5.2.4 Hydrophobic interaction chromatography

The final affinity chromatography technique attempted was hydrophobic interaction chromatography. Three different hydrophobic columns were used: ResourcePhe, ResourceIso and ResourceEth. ResourcePhe is the most hydrophobic column containing phenyl groups, ResourceIso contains isopropyl groups and has intermediate hydrophobicity and the least hydrophobic is ResourceEth which has ether groups. Proteins bind to the column through hydrophobic interactions that are strongest in high salt buffer and are eluted over a concentration gradient of decreasing salt. AlsD bound to the ResourcePhe column but the purity was unchanged after elution. AlsD did not bind to either the Iso or Eth columns and the protein recovered from the flow-through contained the same levels of impurites.

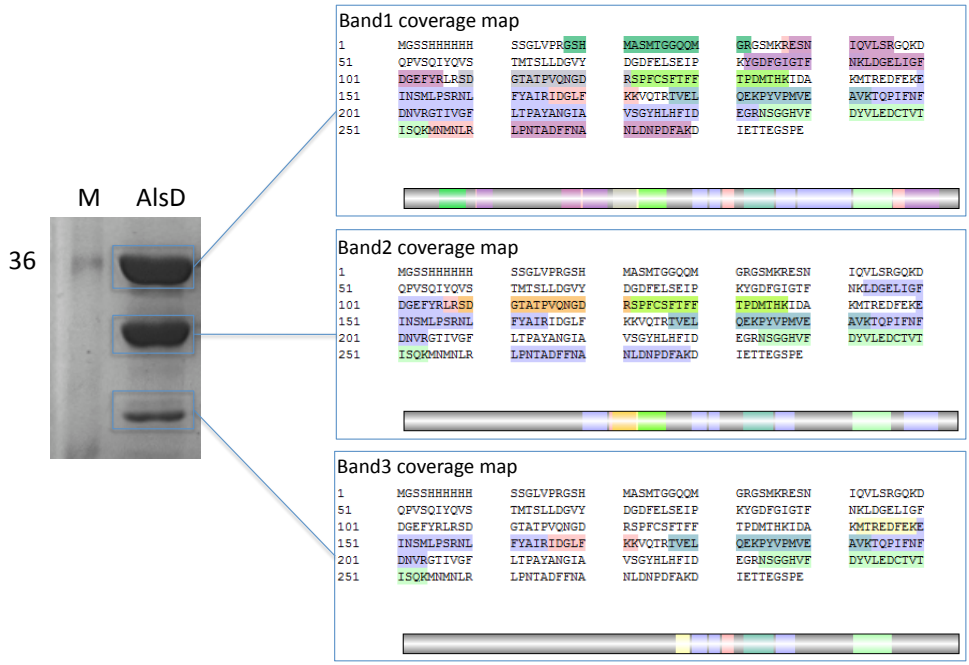
4.5.2.5 Identification of contaminants

As attempts to separate the smaller fragments from the full length AlsD were unsuccessful, a SDS-PAGE gel was submitted for mass spectrometry analysis to identify if the contaminants were C-terminally cleaved degradation products. The three bands highlighted in Figure 4.7 were cut out of the gel and subjected to a tryptic digest. The digested fragments were analysed by mass spectrometry, the coverage maps are also shown in Figure 4.7. All three bands were identified as AlsD but unexpectedly the coverage maps of the tryptic digest fragments appeared to show N-terminal degradation. As the protein has an N-terminal his-tag and the degradation products were still present after IMAC it was expected that the protein was being degraded from the C-terminus. To confirm whether the degradation products were his-tagged a western blot was conducted. The blot (shown in Figure 4.8) confirmed that the full length protein contained a his-tag as expected and that the two truncated degradation products did not. Since these degradation products are not his-tagged but readily co-purify during IMAC, it is likely that they interact with full-length AlsD to form a dimer in which one of the protein chains is truncated. AlsD dimerisation is discussed in Section 1.5.4.

4.5.2.6 Other attempts to generate pure AlsD

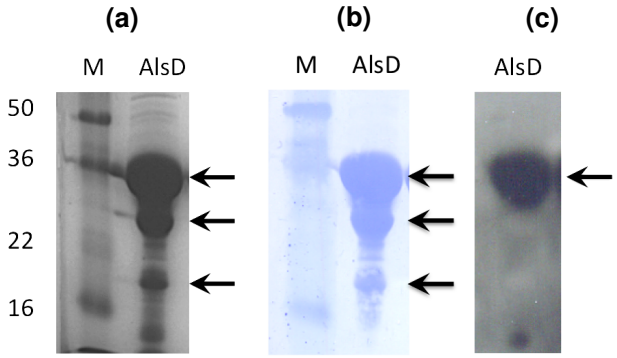
As attempts to separate the full length protein from truncated AlsD were unsuccessful, another strategy was to prevent the proteolytic degradation by adding protease inhibitors during cell lysis and purification. If the protein was being degraded after the cells were lysed, then addition of a cocktail of protease inhibitors may prevent the degradation from occurring. Unfortunately this was unsuccessful and the degradation products were still present regardless of whether protease inhibitors were used. This implied that the degradation was occurring *in vivo* during protein expression rather than after the cells were lysed. Alternatively, process-

Figure 4.7: Protein characterisation by in-gel tryptic digest and mass spectrometry



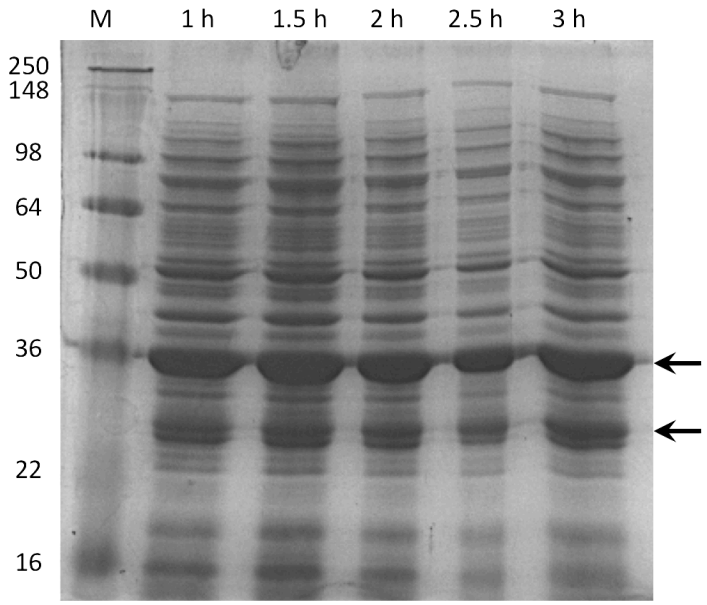
The SDS-PAGE gel pictured was submitted for in-gel tryptic digest and mass spectrometry characterisation. The bands highlighted were cut out of the gel, digested and analysed by mass spectrometry. The coverage maps show all bands were identified as AlsD. For Band1 23 peptides were observed with a 71% coverage, for Band2 13 peptides were observed with 45% coverage and for Band3 8 peptides were observed with 28% coverage.

Figure 4.8: Western blot of *B. subtilis* AlsD



A western blot was conducted to identify if the AlsD degradation products were his-tagged (a) shows an SDS-PAGE gel, M (molecular weight markers), AlsD (IMAC purified AlsD), arrows show the three major protein products, (b) shows the PVDF membrane stained with coomassie, arrows show the three major proteins were transferred to the membrane, (c) shows the western blot, the arrow shows only full length AlsD contains a his-tag.

Figure 4.9: Expression of *B. subtilis* AlsD with 1-3 h induction time at 37°C



The induction time of AlsD expression was investigated. This SDS-PAGE gel shows the total cell extract after a 1, 1.5, 2, 2.5 and 3 h induction time, arrows identify that the full length AlsD and degradation product are detected in all samples.

ing during transcription could result in truncated mRNAs and subsequently truncated protein chains

It was proposed that the protein may be degraded at high yields and that decreasing the induction time might lower the overall yield but lead to a cleaner product. A small scale trial of expression at 37°C for different incubation times was conducted as shown in Figure 4.9. The total soluble fraction after a 1, 1.5, 2, 2.5 and 3 h induction was loaded onto an SDS-PAGE gel. In all samples the over-expression of ALDC was successful but there appears to be no difference in the overall abundance of full length and truncated AlsD.

Taken together, the evidence suggests that truncated AlsD dimerises with full length his-tagged AlsD. As the truncated species proved very difficult to separate from the full-length dimeric species, and because it is a minor contaminant, it was decided to proceed with the

enzyme assays using this material.The final purification protocol was a single IMAC step as described in Section 4.5.2.1.

In the future, for particular experiments, it would be necessary to improve the purity of AlsD. Due to the problems described in this chapter the most successful approach may be to try and find an expression system where this degradation could be avoided. Initially this would involve screening different host cell lines and if this was unsuccessful then re-cloning could be tried. The current vector encodes 23 additional amino acids between the the his-tag and the start of the gene and the N-terminus is flexible as the first 20 residues of *B. brevis* ALDC are not seen in the crystal structure. Clones could be generated without the additional vector encoded sequence and with truncated N-termini as this may protect the recombinant protein against proteolysis.

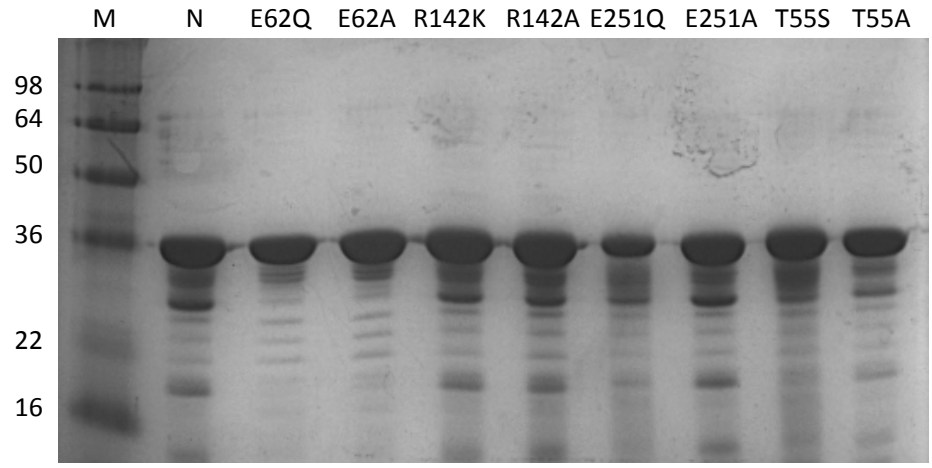
4.6 Expression and purification of AlsD mutants

All eight AlsD mutants were successfully expressed and purified as described for native AlsD. The SDS-PAGE gel in Figure 4.10 shows the final purity of the proteins that were used in enzyme assays. Native, T55S, T55A, R142K, R142A, E251Q and E251A all indicated the presence of truncated AlsD wheres E62Q and E62A did not. In all proteins minor degradation occurred prior to storage at -20 °C but all proteins were stable and underwent no further degradation when stored in aliquots at -20 °C.

4.6.1 Intact mass analysis

Samples of native AlsD and all eight mutants were submitted for intact mass analysis by positive ion ESI mass spectrometry to confirm their identity, as the mutants had a different mass

Figure 4.10: Final purity of native and mutant AlsD



Samples of native (N) and mutant AlsD containing 5 μg of protein were analysed on a 14% SDS PAGE gel. All proteins except E62Q and E62A show evidence of the two contaminating bands identified by mass spectrometry. All samples show some slight degradation occurred during the purification and storage at 4 °C. These proteins were stored in aliquots at -20 °C until required for use in enzyme assays.

Table 4.2: Intact mass analysis of native and mutant AlsD

Protein sample	Calculated average mass	Experimentally measured mass
AlsD	32212.1328	32211.7
AlsD E62Q	32211.1481	32211.1
AlsD E62A	32154.0962	32154.2
AlsD R142K	32184.1194	32184.4
AlsD R142A	32127.0240	32127.2
AlsD E251Q	32211.1481	32211.3
AlsD E251A	32154.0962	32154.4
AlsD T55S	32198.1060	32198.4
AlsD T55A	32182.1066	32181.5

from native AlsD. All proteins were found to have an experimentally determined mass of +/- 0.4 Da (<13 ppm) assuming the N-terminal Met was absent. Table 4.2 shows the predicted average mass for proteins missing the N-terminal Met and the experimentally determined mass for all samples. The mass spectra for all the recombinant proteins are included in Appendix D.

4.6.2 Secondary structure prediction using circular dichroism

To confirm that the mutations had not affected the overall structure of the proteins, native AlsD and all eight mutants were analysed by circular dichroism. The protein concentrations were measured by absorbance at 280 nm and calculated using the extinction coefficient predicted by ProtParam [102]. The protein samples were diluted to a concentration of 0.1 mg/ml in 10 mM phosphate buffer pH 7.6. Far-UV CD spectra were recorded between 190-240 nm, the spectra obtained were in good agreement and were analysed using the CDSSTR algorithm [103–105] and SP175 reference set [82] of the web based server DichroWeb [106–108]. Table 4.3 shows the secondary structure prediction for all AlsD proteins and also *B. brevis* ALDC; the CD spectra are included in Appendix E. The predictions were in good agreement with each other predicting a predominantly beta sheet protein as expected.

4.7 Homology model of AlsD

As the structure of AlsD is unknown, the amino acid sequence was submitted to Phyre2 [83]. Phyre2 is a web server for the prediction of protein structures by homology modelling. The homology modelling was successful, 91% of the sequence was modelled with 100% confidence as containing an alpha-acetolactate decarboxylase fold.

Table 4.3: Secondary structure prediction by dichroweb analysis

Protein sample	α-helix	β-strand	β-turn	Unordered	Total	RMSD
AlsD	0.04	0.39	0.13	0.42	0.98	0.065
AlsD E62Q	0.04	0.39	0.13	0.42	0.98	0.067
AlsD E62A	0.05	0.38	0.13	0.44	1.00	0.069
AlsD R142K	0.05	0.38	0.13	0.42	0.98	0.061
AlsD R142A	0.04	0.41	0.12	0.41	0.98	0.07
AlsD E251Q	0.06	0.37	0.13	0.42	0.98	0.082
AlsD E251A	0.06	0.37	0.13	0.43	0.99	0.065
AlsD T55S	0.05	0.39	0.13	0.42	0.99	0.082
AlsD T55A	0.05	0.38	0.13	0.42	0.98	0.052
B. brevis ALDC	0.07	0.40	0.11	0.41	0.99	0.056

4.8 Conclusion

Two *B. subtilis* genes have been successfully cloned into expression vectors. BdhA was expressed and purified for use in the development of a coupled assay. AlsD was also successfully over-expressed and purified. However it was determined that AlsD suffers from *in vivo* degradation during over-expression and it was not possible to separate the truncated protein chains, most probably because they form dimeric assemblies with full-length AlsD chains. Full-length/full-length AlsD dimer is still the predominant species, with the mixed full-length/truncated AlsD dimer a minor component, so the material was deemed suitable for use in further experiments.

Eight active site mutants of AlsD were successfully engineered, expressed and purified. Mutant plasmids were confirmed by DNA sequencing, and mutant proteins were confirmed by intact mass spectrometry and CD analysis confirmed that all the proteins had a similar proportion of secondary structure and none of the mutants had a vastly different structure.

The secondary structure prediction was also in very good agreement with *B. brevis* ALDC which was of high purity, further suggesting that the degradation products were correctly folded proteins.

Chapter 5

Assay development and kinetics of ALDC

5.1 Introduction

The kinetics of an enzyme catalysed reaction can provide information about the enzyme mechanism and the effectiveness of inhibitors while assays of active site mutants can indicate the important residues for catalysis. To study enzyme kinetics it is necessary to have an accurate and quantitative assay. In a direct assay the rate of reaction is determined by measuring the disappearance of substrate or the appearance of product. If it is not possible to measure substrate or product, an assay can be developed which allows indirect measurement of the product by coupling the product to a second reaction.

5.2 Aims

5.2.1 Development of an ALDC activity assay

The first aim of the work described in this chapter is to develop a quantitative assay for the activity of ALDC. Previous biochemical studies of ALDC have used a modified Voges-Proskauer assay to indirectly detect the presence of acetoin. This method has several limitations; the colourimetric reaction is non-specific requiring many control reactions and as a stopped assay the rate is not directly measured. Improvements to the VP method will be sought and alternative continuous assays developed.

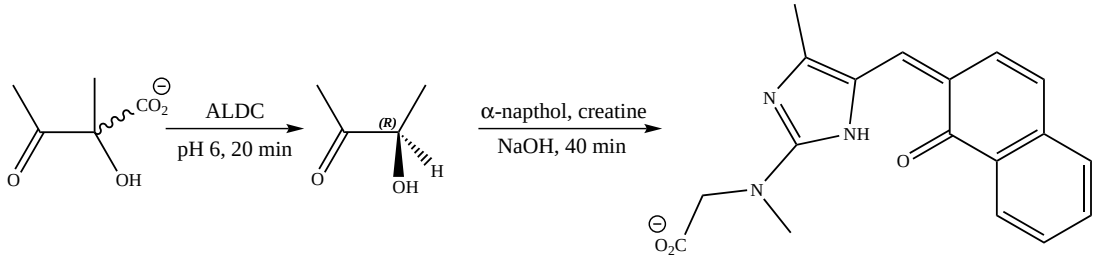
5.2.2 Analysis of the kinetics of active site mutants

To gain insight into the molecular mechanism of ALDC, the activity of wild-type and active site mutants will be assessed. The mutants' activity should indicate which amino acid residues are important in catalysis. Any mutants that show activity will be fully characterised to compare their kinetic parameters to wild-type AlsD.

5.2.3 Analysis of the inhibition of ALDC by transition state analogues

Transition state analogues developed in Chapter 3 will be assayed to determine the inhibition constants (K_i). Enzymes will be assayed in the presence of increasing concentrations of inhibitor so that kinetic parameters and mode of inhibition can be determined by global non-linear regression, to identify which transition state analogues are suitable targets for X-ray crystallography co-crystallisation studies.

Figure 5.1: Voges-Proskauer assay for the detection of acetoin



3 (*R*)-**4** **20**

*The Voges-Proskauer assay is a stopped assay to detect acetoin. Acetolactate **3** is incubated with ALDC at pH 6. The reaction is stopped after 20 min by the addition of a colour reagent which contains α-napthol and creatine dissolved in NaOH. The reaction mixture is incubated for 40 min to allow coloured complex **20** to form [113,114]. The absorbance of **20** can be monitored at 522 nm.*

5.3 Development of the Voges-Proskauer test as an ALDC activity assay

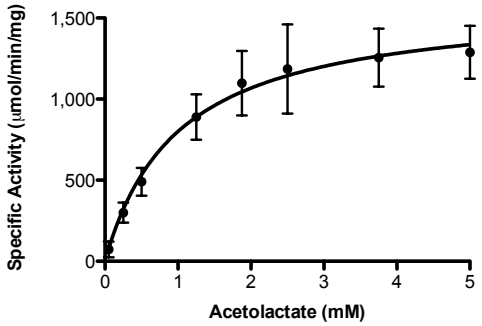
The activity of *Bacillus brevis* ALDC (a gift from Novozymes) was studied using the Voges-Proskauer (VP) assay. The assay reaction scheme is shown in Figure 5.1; it is a stopped assay for the detection of acetoin.

This assay is based on the original observation of Voges and Proskauer [109] that certain microbial glucose peptone cultures underwent a colour change upon the addition of a KOH solution. Subsequent studies [110,111] indicated that the colour results from the reaction of acetoin with a compound containing a guanidine group and that the colour was intensified by the addition of α-naphthol [112]. The assay conditions used in this study were based on the Sigma-Aldrich protocol (Appendix C), ALDC was incubated with racemic acetolactate for 20 min then the reaction was stopped by the addition of a colour reagent containing creatine and α-naphthol in 1M NaOH. The mixture was incubated for 40 min to allow the formation of the coloured complex shown in Figure 5.1.

Similar versions of the VP assay have been used in previous studies of ALDC to give information about the kinetics of the reaction [41,46,115]. It also has applications in microbiology as part of IMViC (indole, methyl red, Voges-Proskauer and citrate) tests which is a combination of four tests that are used for the identification of enteric bacteria [116]. A modified VP assay is one of a number of techniques used by the brewing industry to determine beer diacetyl and acetoin concentrations [117].

It was challenging to get reproducible results using this assay and several problems were identified. It was necessary to prepare the enzyme solution just before use and use a fresh aliquot of enzyme for each assay as activity was lost if the enzyme was freeze-thawed. The enzyme substrate is unstable and can spontaneously decarboxylate so it had to be prepared immediately before use and the assay was conducted with controls for each substrate concentration to account for this. It was also identified that the colour reagent is unstable, as inconsistent results were obtained depending on how fresh the solution was. To overcome this issue the colour reagent was prepared just before use. With all these challenges identified and corrected for as described, Figure 5.2 shows the results from three replicates. It is evident that there is still a lot of variation using this assay*.

Figure 5.2: Kinetics of *B. brevis* ALDC measured by the Voges-Proskauer assay



The mean of three replicates are plotted. Errors bars show one standard deviation above and below the mean. Prism was used for non-linear regression curve fitting to the Michaelis-Menten equation.

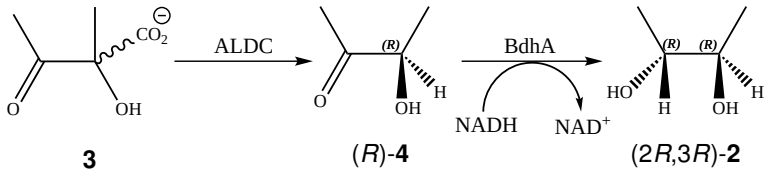
*Preliminary studies involving the VP assay were conducted by the author during her MSc research [118].

A stopped assay is not ideal for studying enzyme kinetics, as it is not possible to directly measure the rate in real time. Instead, the amount of product produced is measured and a rate is calculated based on the incubation time. It is not always possible to confirm that the measured initial rate is linear over time. Due to the practical difficulties and problems with a stopped assay it was decided to discontinue the work on the optimisation of the Voges-Proskauer assay and to develop an alternative, more quantitative assay.

5.4 Development of a coupled ALDC assay using BdhA

A continuous assay for the study of ALDC has been developed by using the coupling enzyme BdhA. BdhA is the downstream enzyme of ALDC in the butanediol fermentation pathway which reduces (*R*)-acetoin (*R*)-**4** to (2*R*,3*R*)-butanediol (2*R*,3*R*)-**2** using the cofactor NADH. The assay scheme is shown in Figure 5.3. BdhA and NADH are added to the reaction mixture so that as ALDC produces (*R*)-acetoin (*R*)-**4** it is reduced by BdhA to (2*R*,3*R*)-butanediol (2*R*,3*R*)-**2**. For every molecule of acetoin reduced one molecule of NADH is oxidised to NAD^+ and this can be monitored spectroscopically at 340nm. Monitoring the loss of signal from NADH at 340nm is therefore proportional to the turnover of the enzyme.

Figure 5.3: Coupled assay for the activity of ALDC



3 (*R*)-**4** (2*R*,3*R*)-**2**

*A coupled assay to study ALDC kinetics was developed. Acetolactate **3** is incubated with ALDC in the presence of BdhA and NADH. BdhA is present in excess so that as ALDC turns over acetolactate **3** to (R)-acetoin (R)-**4** BdhA converts it to butane-2,3-diol (2R,3R)-**2**. In this process NADH is oxidised to NAD^+. NADH is UV active and the loss can be measured at 340 nm.*

5.4.1 Activity of BdhA

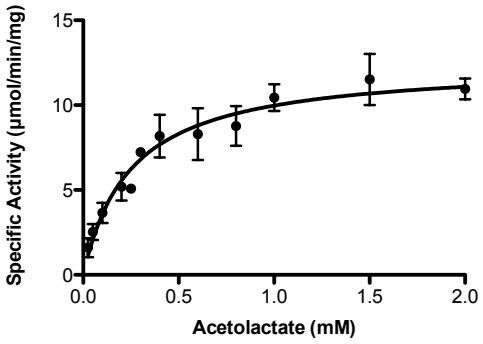
The expression and purification of BdhA was described in Chapter 4 but before it could be used in a coupled assay the activity had to be assessed. A previous study [53] reported the activity of BdhA and its specificity to reduce (*R*)-acetoin (*R*)-**4** to give (2*R*,3*R*)-butanediol (2*R*,3*R*)-**2**. The activity of BdhA expressed in this study was tested under the conditions previously described [53]. Under these conditions the protein showed a similar K_{m} to the literature value, but k_{cat} was lower. As the pH optimum of ALDC is lower, the activity of BdhA was tested at a lower pH. The VP assay was conducted at pH 6.0, however BdhA was unstable at this pH and precipitated out of solution. The activity of BdhA was assayed in bis-tris buffer at pH 6.5. The enzyme was more active at this pH, the K_{m} dropped to 0.1 mM and the k_{cat} increased. The kinetic parameters are outlined in Table 5.1: although the BdhA activity was lower than the value reported in the literature the enzyme activity was sufficient for use in the enzyme assay as it was used in excess.

Table 5.1: Kinetic parameters of BdhA

pH	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})	**Reference**
6.5	0.11 ± 0.13	0.064 ± 0.012	0.58	This study
8.0	0.20 ± 0.09	0.022 ± 0.002	0.11	This study
8.0	0.26	0.98	3.8	[53]

The activity of BdhA was measured at two different pH (this study), the data was analysed in Prism and the kinetic parameters summarised and compared to a previous study [53].

Figure 5.4: Kinetics of AlsD measured by the coupled assay



AlsD activity was measured using the coupled assay. The means of three replicates are plotted; error bars show one standard deviation above and below the mean. Curve fitting to the Michaelis-Menten equation by non-linear regression was carried out using Prism.

5.4.2 ALDC activity measured by the coupled assay

Figure 5.4 shows results obtained from the coupled assay for AlsD. The assay was able to successfully measure the activity of both *B. subtilis* AlsD (triplicate) and *B. brevis* ALDC (single replicate) and the kinetic parameters are summarised in Table 5.2. Some problems were encountered when trying to measure enzyme inhibition using this assay. The ALDC inhibitor was found to inhibit the coupling enzyme and cause the coupling enzyme to precipitate out of solution. Due to the problems trying to measure inhibition using the coupled assay it was decided to explore an alternative method to assay ALDC inhibition.

5.5 Development of a circular dichroism ALDC assay

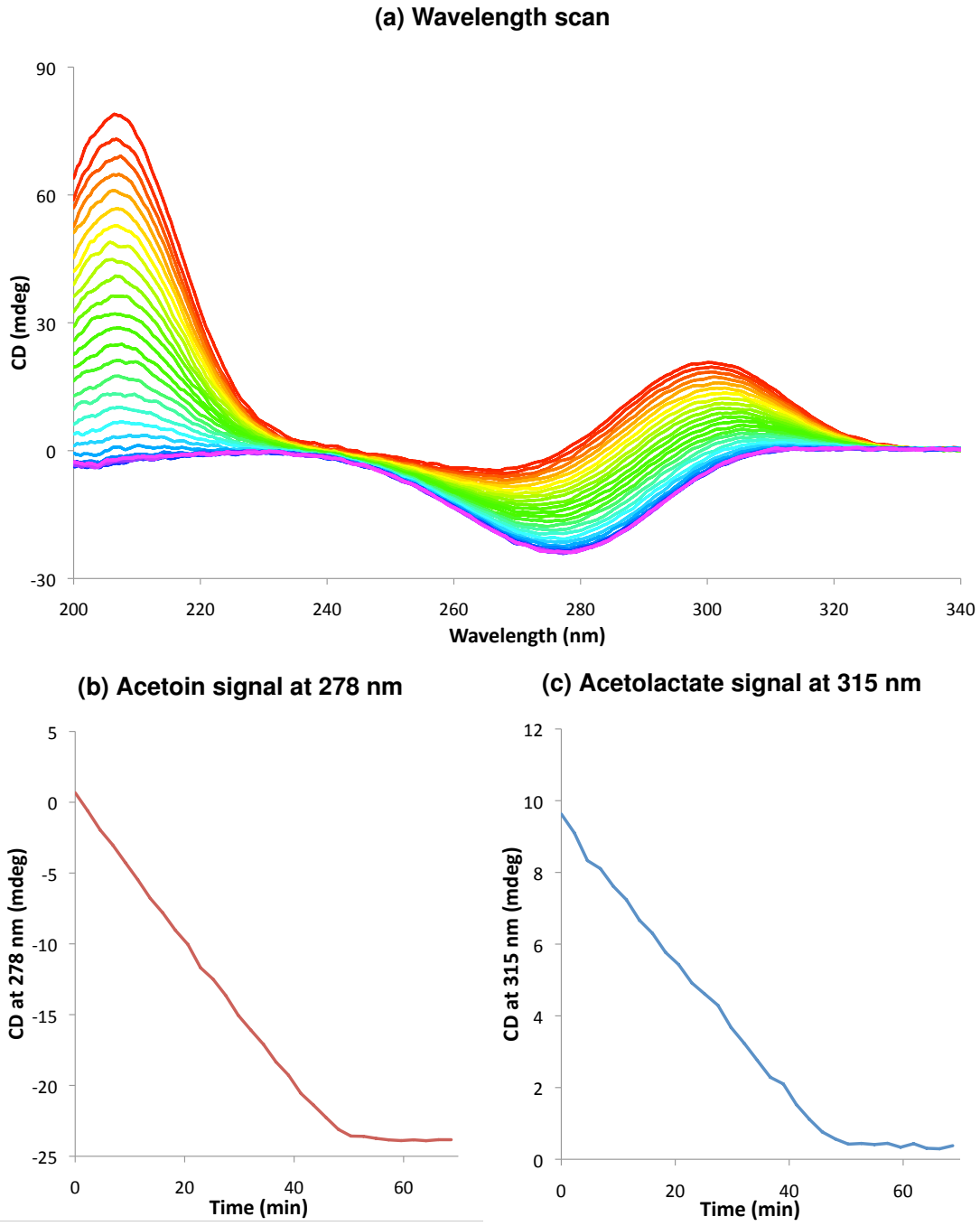
Due to the inherent chirality of the ALDC catalysed reaction it was identified that it would be possible to study ALDC activity using circular dichroism. A literature search identified that Vinogradov et al. [95] had described monitoring the activity of ALDC using CD in the

development of a coupled assay for acetolactate synthase.

5.5.1 CD spectra of substrate and product

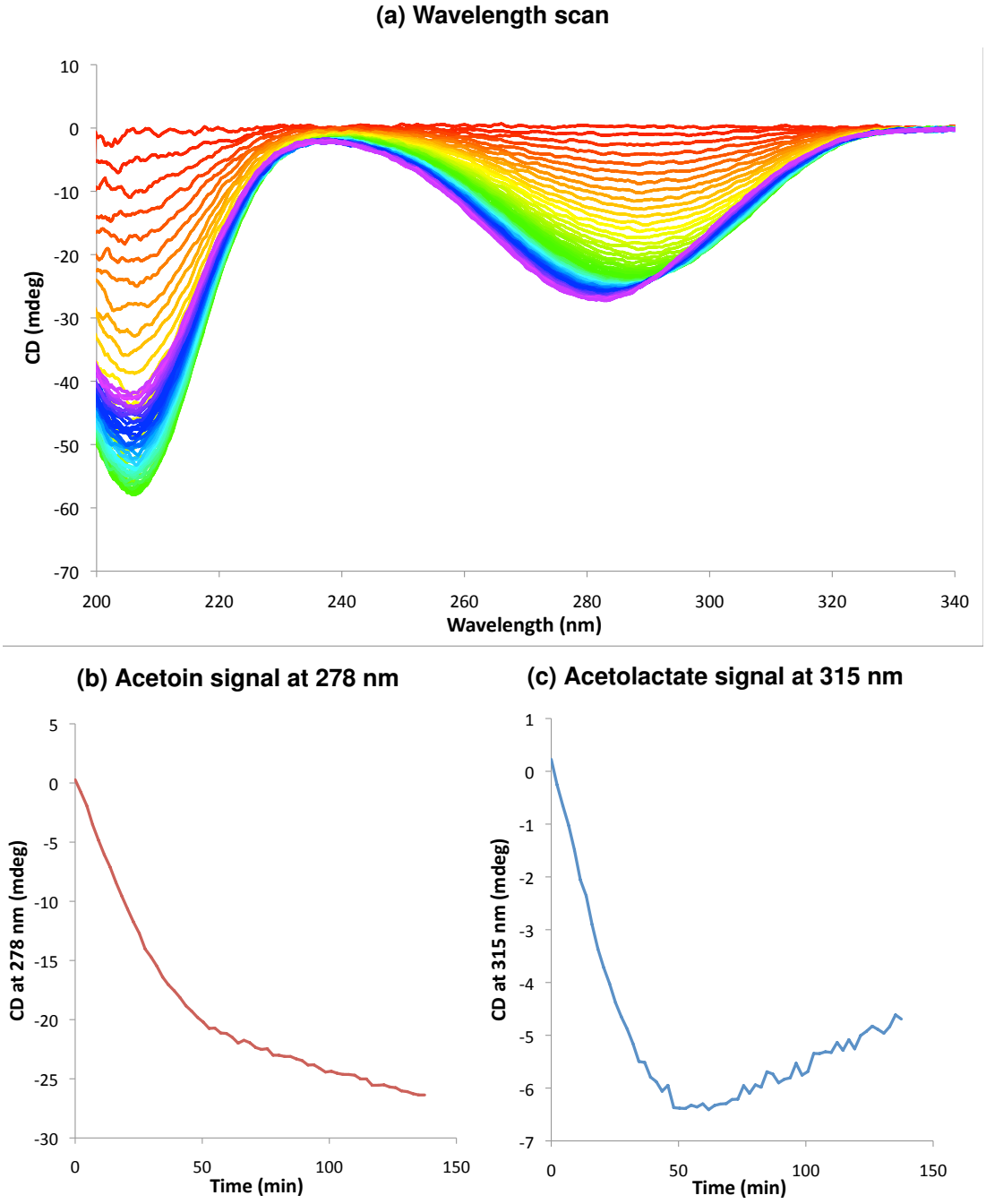
The CD spectra of (*S*)-acetolactate (*S*)-**3** and (*R*)-acetoin (*R*)-**4** are reported [95] and are different enough to allow for the monitoring of enzyme activity using CD. To confirm that this assay would be suitable to study the kinetics of ALDC a continuous wavelength scan was conducted. Figure 5.5a shows the *B. brevis* ALDC catalysed reaction of (*S*)-acetolactate (*S*)-**3** into (*R*)-acetoin (*R*)-**4**. The differences between the spectra make it possible to choose a wavelength where either the disappearance of substrate, or the appearance of product, can be monitored. The two wavelengths that were identified were 278 nm and 315 nm, the CD signal at these two wavelengths are shown in Figure 5.5b and 5.5c respectively. At 278 nm the formation of (*R*)-**4** can be monitored, this is the point where the CD signal for (*S*)-**3** crosses the axis and has no signal. As it is a negative peak the slope is negative but is proportional to the rate of formation of (*R*)-**4**. At 315 nm the CD signal for (*R*)-**4** is 0 and only the tail of a positive (*S*)-**3** peak is observed. This is the disappearance of a positive peak so the rate observed is the loss of (*S*)-**3** over time. Figure 5.6a shows a wavelength scan using racemic acetolactate ($\pm$)-**3** as ALDC substrate. Initially there is no CD signal as the substrate is racemic but as the enzyme is added the spectrum for (*R*)-**4** and (*R*)-**3** are formed. The signal for (*R*)-**3** is observed because as the enzyme turns over (*S*)-**3** it leaves an excess of (*R*)-**3**. At 278 nm (Figure 5.6b) the formation of (*R*)-**4** can be observed initially at a faster rate followed by a slower rate. At 315 nm (Figure 5.6c) the **3** signal initially begins at 0 and is followed by the fast formation of a negative peak due to an excess of (*R*)-**3** as the enzyme turns over (*S*)-**3** and then the peak begins to disappear at a slower rate as the enzyme turns over the remaining (*R*)-**3**.

Figure 5.5: ALDC catalysed turnover of (*S*)-acetolactate monitored by CD



*The ALDC catalysed reaction of (S)-acetolactate (S)-**3** was monitored by continuous wavelength scan CD, (a) the entire spectra are shown starting from (S)-acetolactate (S)-**3** (blue) to (R)-acetoin (R)-**4** (purple). The reaction went to completion in approximately 50 min but was monitored for 68.7 min. (b) At 278 nm there is no signal for (S)-acetolactate (S)-**3** and the change in CD signal is due to the formation of (R)-acetoin (R)-**4**. (c) At 315 nm there is no signal for (R)-acetoin (R)-**4** and the change in CD signal is due to the disappearance of (S)-acetolactate (S)-**3**.*

Figure 5.6: ALDC catalysed turnover of (±)-acetolactate monitored by CD



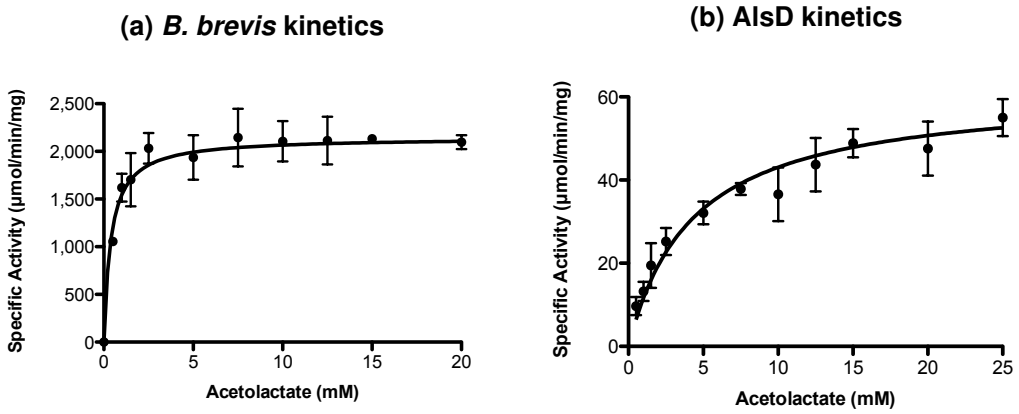
*The ALDC catalysed reaction of (±)-acetolactate (±)-**3** was monitored by continuous wavelength scan CD, (a) the entire spectra are shown starting from (±)-acetolactate (±)-**3** (blue) to a mixture of unreacted substrate (R)-acetolactate (R)-**3** and product (R)-acetoin (R)-**4** (purple). The reaction was monitored for 137.5 min and did not reach completion. (b) At 278 nm there is no signal for acetolactate **3** and the change in CD signal is due to the formation of (R)-acetoin (R)-**4**. (c) At 315 nm there is no signal for (R)-acetoin (R)-**4** and the change in CD signal is due to the appearance of a peak for (R)-acetolactate (R)-**3** followed by the disappearance of this peak.*

These two experiments demonstrated that it would be possible to study the ALDC catalysed reaction by monitoring the CD signal of either substrate or product. As the signal for product was stronger than for substrate it was decided to monitor assay reaction rates at 278 nm. The rate was converted from units of mdeg/s into units of mM/s by using molar ellipticity. The rates were then converted into specific activity with units of micromol/mg/min to allow comparison between different assay techniques. The conversion of units are shown in Appendix F.

5.5.2 Activity and kinetics of *B. subtilis* recombinant wild-type AlsD and mutants

Figure 5.7b shows the activity of *Bacillus subtilis* AlsD that was purified as described in Chapter 4. The activity was studied over a concentration range of 0.5-25 mM (*S*)-acetolactate (*S*)-**3**, as the strength of the CD signal limited the lower substrate concentration to 0.5 mM. Non-linear curve fitting in Prism was conducted on 3 replicates and the kinetic parameters are summarised in Table 5.2. The activity of the eight active site mutants described in Chapter 4 were assayed in triplicate at a single substrate concentration, as shown in Figure 5.8a. Five of the mutants (E62Q, E62A, R142A, E251Q and E251A) showed very little activity; the rates were too low to be accurately measured. Three of the mutants were active; both threonine mutants T55S and T55A showed higher activity than native AlsD. T55S was approximately 2 fold more active and T55A was approximately 2.5 fold more active. When arginine (R142) was mutated to lysine the activity was 40 % of native AlsD. The three active mutants were studied in more detail to measure kinetic parameters, as shown in Figure 5.8b. The kinetic parameters are summarised in Table 5.2. T55A and T55S have a similar K_m to native AlsD but have an increased k_{cat}, R142K has a decreased K_{m} and a decreased k_{cat}. The decreased K_{m} implies that the lysine mutant binds substrate with a greater affinity but

Figure 5.7: Kinetics of *B brevis* ALDC and AlsD measured by CD



The activity of B. brevis ALDC and B. subtilis AlsD were measured over a substrate concentration range of 0.5-20 mM and 0.5-25 mM respectively. Enzyme activity was measured by observing the change in CD at 278 nm and converted into units of specific activity (µmol of acetoin produced min^{-1} mg of protein^{-1}). The mean of three replicates are plotted and errors bars show one standard deviation above and below the mean. Curve fitting was conducted using Prism and the kinetic parameters are summarised in Table 5.2.

it is less effective as a catalyst and turns over substrate much slower.

These results imply that both the glutamate residues are essential for catalysis. Lysine can partially compensate for the loss of the arginine residue, implying a basic residue is necessary for catalysis. As the threonine mutants were more active than native AlsD it can be concluded that this residue is not involved in the catalytic mechanism of ALDC.

5.5.3 Kinetics and inhibition of *B. brevis* ALDC

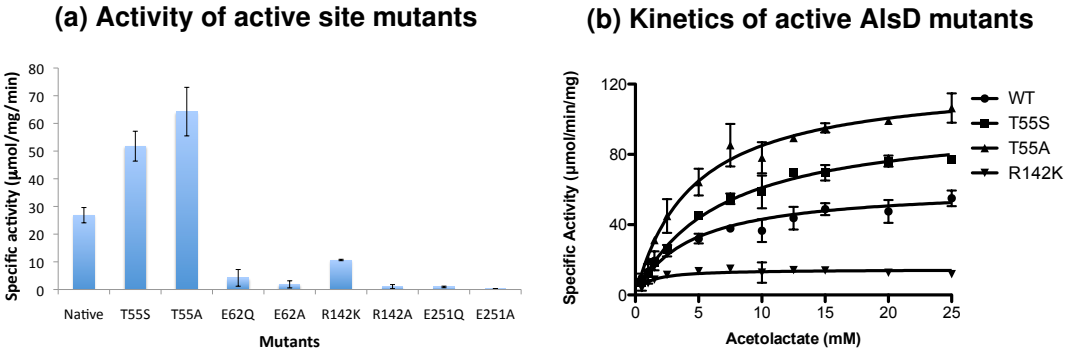
The activity of *B. brevis* ALDC was also studied using the CD assay and the curve fitting is shown in Figure 5.7a. The kinetics of *Bacillus brevis* ALDC were calculated based on curve fitting to three replicates. The kinetic parameters are summarised in Table 5.2. *B. brevis* ALDC has a lower K_{m} and greater k_{cat} than *B. subtilis* ALDC. The enantiomerically enriched 2,3-dihydroxy-2-methylbutanoic acids **11** that were synthesised as described in

Table 5.2: Kinetics parameters of *Bacillus subtilis* and *Bacillus brevis* ALDC measured by all 3 assays

Assay	Organism	Enzyme	Substrate	Buffer	pH	V_{max} (μmol min^{-1} mg^{-1})	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})	R^2
VP	*B. brevis*	Wild-type	(±)-**3**	MES	6.0	1604 ± 243	1.01 ± 0.47	759.7 ± 114.9	755.2	0.91
NADH	*B. subtilis*	Wild-type	(±)-**3**	Bis-Tris	6.5	12.47 ± 1.23	0.25 ± 0.08	5.99 ± 0.59	23.94	0.92
NADH	*B. brevis*	Wild-type	(±)-**3**	Bis-Tris	6.5	137.7 ± 14.9	0.09 ± 0.04	66.71 ± 7.22	717.0	0.95
CD	*B. brevis*	Wild-type	(*S*)-**3**	$NaPO_4$	6.0	2147 ± 86	0.39 ± 0.13	1040 ± 42	2657	0.92
CD	*B. subtilis*	Wild-type	(*S*)-**3**	$NaPO_4$	6.0	61.64 ± 6.17	4.31 ± 1.47	29.59 ± 2.96	6.87	0.90
CD	*B. subtilis*	T55A	(*S*)-**3**	$NaPO_4$	6.0	123.2 ± 9.3	4.50 ± 1.15	59.13 ± 4.46	13.14	0.97
CD	*B. subtilis*	T55S	(*S*)-**3**	$NaPO_4$	6.0	100.8 ± 9.1	6.53 ± 1.54	48.38 ± 4.37	7.41	0.98
CD	*B. subtilis*	R142K	(*S*)-**3**	$NaPO_4$	6.0	14.53 ± 1.31	0.49 ± 0.75	6.97 ± 0.63	7.24	0.96

The kinetic parameters from all three assays are summarised, all the curve fitting was carried out using Prism and had a R^2 value of $\geq$ 0.9. All the assays were carried out under different conditions as described.

Figure 5.8: Activity and kinetics of AlsD mutants

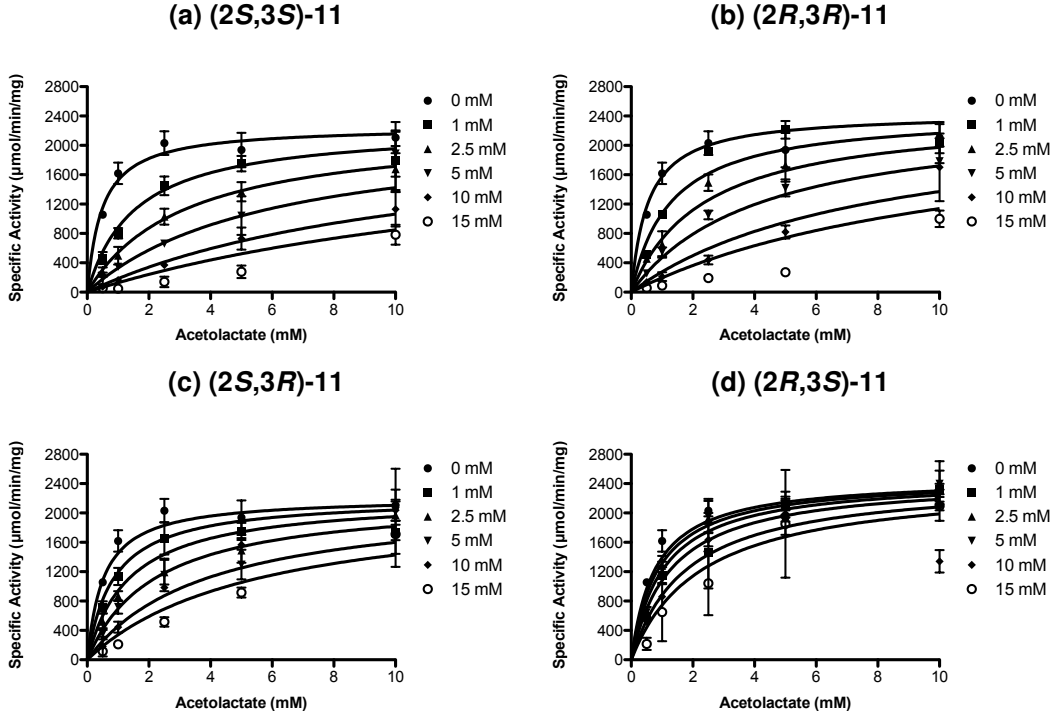


(a) The activity of native AlsD and eight active site mutants were tested in triplicate, the mean is plotted and the error bars show one standard deviation above and below the mean. (b) The mutants that showed activity (T55S, T55A, R142K) were studied in more detail. The kinetics of the three active mutants were tested under the same conditions described for native AlsD. Measurements were obtained in duplicate and are plotted with native AlsD, curve fitting to the Michaleis-Menten equation was conducted in Prism. The kinetic parameters are summarised in Table 5.2.

Chapter 3 were tested as inhibitors of *B. brevis* ALDC. It was decided to test the inhibition on *B. brevis* ALDC as this was the protein that was used in crystallisation studies. The assay was conducted using 5 substrate concentrations from 0.5-10 mM and five inhibitor concentrations from 1-15 mM. The resulting rate data was subjected to global curve fitting using Prism to determine the mode of inhibition and the K_i. The plots are summarised in Figure 5.9 and the kinetic parameters in Table 5.3. Three of the inhibitors were identified as competitive inhibitors, (2*S*,3*S*)-**11**, (2*R*,3*R*)-**11** and (2*S*,3*R*)-**11** and the final compound (2*R*,3*S*)-**11** was identified as a mixed inhibitor.

These results imply that the three 2,3-dihydroxy-2-methylbutanoic acids that act as competitive inhibitors would be good targets for co-crystallisation studies with ALDC and it shows that both (*S*)- and (*R*)-acetolactate transition state analogues inhibit ALDC.

Figure 5.9: Inhibition of *B brevis* ALDC by chiral 2,3-dihydroxy-2-methylbutanoic acids



*The four chiral 2,3-dihydroxy-2-methylbutanoic acids **11** (Chapter 3) were tested as inhibitors of B. brevis ALDC. Enzyme activity was measured at five substrate concentrations (0.5-10 mM) and five inhibitor concentrations (1-15 mM) for each inhibitor in duplicate. The mode of inhibition and kinetic parameters (Table 5.3) were determined by global curve fitting in Prism.*

Table 5.3: Inhibition parameters

Inhibitor	Mode of inhibition	V_{max} (μmol min^{-1} mg^{-1})	K_m (mM)	K_i (mM)	R^2
(2*S*,3*S*)-**11**	Competitive	2269 $\pm$ 177	0.49 $\pm$ 0.14	0.46 $\pm$ 0.18	0.97
(2*R*,3*R*)-**11**	Competitive	2440 $\pm$ 247	0.55 $\pm$ 0.26	0.76 $\pm$ 0.40	0.94
(2*S*,3*R*)-**11**	Competitive	2216 $\pm$ 209	0.56 $\pm$ 0.28	1.72 $\pm$ 0.98	0.92
(2*R*,3*S*)-**11**	Mixed	2482 $\pm$ 320	0.78 $\pm$ 0.43	7.68 $\pm$ 8.89	0.88

*Enantiomerically enriched 2,3-dihydroxy-2-methylbutanoic acids **11** were tested as inhibitors of B. brevis ALDC. The kinetic parameters are summarised, (2S,3S)-**11**, (2R,3R)-**11** and (2S,3R)-**11** are competitive inhibitors and (2R,3S)-**11** is a mixed inhibitor.*

5.6 Conclusions

The VP assay for measuring ALDC activity was found to generate inconsistent and irreproducible results, therefore two alternative, more quantitative assays were developed. The first of these alternative assays, the coupled assay, is much more suitable for enzyme kinetic studies. It is easier to carry out than the VP assay, and is a continuous assay, but has added complexity due to the addition of the coupling enzyme and NADH into the reaction mixture. The high extinction coefficient of NADH means this assay is very sensitive and low substrate concentrations can be used. As it is a continuous assay it is more suitable than the VP assay to study initial rates for enzyme kinetic studies. Although it has not been carried out in this study, as this is a spectroscopic assay it would also be possible to transfer to a plate reader spectrophotometer for higher throughput of samples and quicker analysis. However the major stumbling block with this assay is with inhibition studies. The substrates of both ALDC and BdhA are very similar, and ALDC inhibitors also inhibit BdhA. The CD assay overcomes the problems of inhibition of the coupling enzyme in the coupled assay. The CD assay is a simpler system that allows the direct measurement of either substrate or product in a continuous assay but it does require higher concentrations of substrate to detect the signal, but this is a price worth paying for the quality of the data that the assay can yield. In this work CD assay measurements were conducted one at a time in a 1 mm pathlength cuvette to keep reaction volumes to a minimum. This ensured that all reactions could be carried out on a single batch of substrate to avoid potential differences due to substrate purity and enantiomeric excess. It would be possible to carry out these assays using a six sample changer if larger reaction volumes were used. It is not possible to directly compare all three assays as they were individually optimised and carried out under different conditions. Future work would be to carry out all assays under the same conditions to allow direct comparisons.

Chapter 6

X-ray crystallography of ALDC in complex with transition state analogues

6.1 Introduction

X-ray crystallography is a powerful tool in the elucidation of enzyme mechanisms as it provides accurate information about the location of catalytic residues in the active site. Crystal structures of an enzyme in complex with an inhibitor that mimics steps along the reaction co-ordinate provide convincing evidence for a mechanism [119]. For example, crystal structures of isopenicillin N synthase with bound substrate and trapped in a intermediate reaction state allowed the mechanism to be elucidated [120,121]. A further example is that of acetoacetate decarboxylase; the crystal structure of AADase in complex with acetoacetone identified that electrostatic repulsion by an active site glutamate and stabilisation by an arginine residue are important for catalysis [26]. The structure of ALDC in complex with product analogue ethylene glycol identified the location of the active site and revealed that metal coordination is likely to be important in substrate binding. However, these data provided no further insight

into the interaction of either (*S*)- or (*R*)-acetolactate with the ALDC active site.

6.2 Aims

6.2.1 Generate crystal structures of ALDC in complex with transition state analogues

The aims of the work described in this chapter are to optimise conditions for the growth of ALDC crystals and to co-crystallise ALDC in the presence of the chiral transition state analogues that were synthesised in Chapter 3, in order to derive high resolution structures. These structures will provide insight into the mode of binding of (*S*)- and (*R*)-acetolactate.

6.3 Reproducing published crystallisation conditions

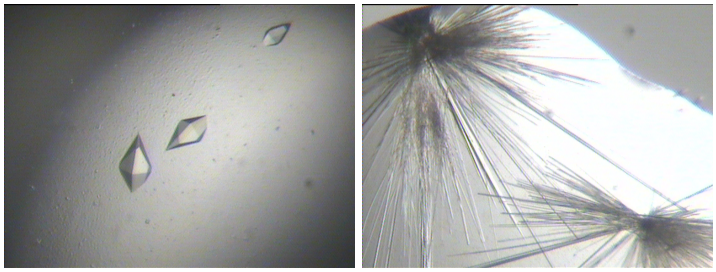
The structure of *Bacillus brevis* ALDC was solved by the Fülöp group. A preliminary publication [63] described three crystallisation conditions and the data collection and processing statistics. The structure was subsequently solved by single-wavelength anomalous diffraction (SAD) on the active site Zn^{2+} and some bound Cd^{2+} from the crystallisation mother liquor [64]. Initially the aim of this work was to reproduce the conditions previously described. The protein used in both cases was from the same batch as it had been supplied as a gift from Novozymes and stored in the -80°C freezer.

6.3.1 Type 1 crystals

The first crystal type (called Type 1) described by Najmudin et al. [63] were bipyramidal crystals belonging to space group $P4_12_12$. These crystals were reproduced under the same conditions as previously reported, as shown in Figure 6.1a. Some of the crystals were cryo-cooled and taken to the European Synchrotron Radiation Facility (ESRF). Datasets were collected on 4 crystals* up to a resolution of 2.0 Å. This was initially very promising but future attempts to reproduce these diffraction quality crystals were unsuccessful. It was possible to grow the crystals under the same conditions but the diffraction was considerably worse. To determine whether the problem was due to the poor cryoprotection, crystals were capillary mounted but a room temperature collection also showed poor diffraction. The problems with these crystals were never solved and work moved on to other conditions.

Figure 6.1: Reproducing previous ALDC crystals

(a) Type 1 **(b) Type 2**



Crystallisation took place in 24 well plates using the hanging drop vapour diffusion method. (a) Type 1 bipyramidal crystals were grown in a mother liquor containing 16% PEG 8K, 80 mM MES pH 6.5 and 200 mM $Zn(OAc)_2$ (b) Type 2 rectangular crystals were not observed, instead needle clusters formed when grown in mother liquor containing 15% PEG 2K MME, 0.1 M Tris pH 8.8, 5 mM $CdCl_2$ and 15% ethylene glycol.

*These crystals were prepared by the author during her MSc research [118], whereas subsequent data collection and further attempts at optimisation were carried out during the PhD.

6.3.2 Type 2 crystals

A second crystal type (Type 2) described by Najmudin et al. [63] were rectangular crystals belonging to space group $P3_221$. It was reported that these crystals grew in 20% polyethylene glycol 2000 monomethyl ether (PEG 2K MME), 0.1 M Tris-HCl pH 7.0, 10 mM $CdCl_2$. A screen was conducted around these conditions, varying the pH 6.0-8.5 and the PEG concentration of 12.5-30%. At pH 8.0-8.5 long needle clusters were observed as shown in Figure 6.1b but the rectangular crystals could not be reproduced. After a discussion with Dr Najmudin who carried out the original work it emerged that further optimisation had been conducted after the publication and the conditions were varied to pH 9.0 with the addition of 15% ethylene glycol. Screening around these further optimised conditions was carried out but the rectangular crystals described by Dr Najmudin were not observed.

6.4 Screening and developing new crystallisation conditions

As attempts to reproduce Type 1 and Type 2 ALDC crystals had been unsuccessful it was decided to see if alternative conditions could be found using commercial screens.

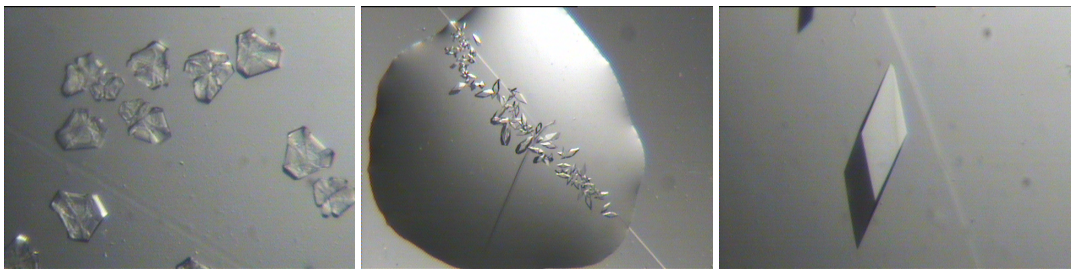
6.4.1 Commercial screening and optimisation

Two 96 well commercial screens (JCSG-*plus* and PACT *premier*) were set up using the Honeybee 963 crystallisation robot. The screens were monitored and after approximately 3 months some promising hits were identified from both screens. The PACT *premier* screen identified hits in A1-3 and B2-4, H3 and H10 and the JCSG-*plus* screen had a single hit in

G4. Four 24 well screens were set up to screen around the hits identified. Nothing positive came from the PACT *premier* hits however after approximately 3 months some crystalline material was observed in the JSCG-*plus* screen as can be shown in Figure 6.2a. These crystals were not suitable for diffraction studies as they are not single crystals. One of the challenges of protein crystallisation is finding conditions that are optimal for both stages of crystallisation; nucleation and growth. In this case the concentration required for nucleation was higher than the optimum for crystal growth resulting in crystals that had grown too rapidly in multiple directions. It was decided that these crystals might be suitable for microseeding. Microseeding uses crystals that have already nucleated as seeds for crystal growth in freshly equilibrated drops [122]. This allows the crystals to be grown under more controlled conditions. The crystals shown in Figure 6.2a were crushed with an acupuncture needle to create a seed stock. Microseeds were transferred to fresh drops by loading an acupuncture needle with microseeds and streaking it through a freshly equilibrated drop [123]. The seeding process was successful and Figure 6.2b shows crystals growing along a seed streak and Figure 6.2c shows a single crystal.

Figure 6.2: Optimising crystallisation conditions

(a) Original conditions **(b) Streak seeding** **(c) Single crystal**



(a) These crystals were grown using the hanging drop vapour diffusion method in a mother liquor containing 30% PEG 2K MME, 0.1 M Tris pH 8.5 and 0.2 M TMAO, they appeared after 3 months (b) these crystals were grown using the streak seeding method, the crystal tray was prepared and equilibrated overnight and streak seeded with an acupuncture needle (c) a large single crystal grown by streak seeding into a mother liquor containing 24% PEG 2K MME, 0.1 M Tris pH 8.0 and 0.2 M TMAO.

These crystals belong to space group $P3_221$ which is the same as the previously reported [63] type 2 crystals. They were also grown under very similar conditions, with type 2 crystals grown in 20% PEG 2K MME, 0.1 M Tris pH 9.0, 10 mM $CdCl_2$ and 15% EG (Dr Najmudin personal communication) and these crystals grew in 28-33% PEG 2K MME, 0.1 M Tris pH 8.5 and 0.2 M triethylamine N-oxide (TMAO). The crystals grew very reproducibly and very quickly; crystals were usually visible a few hours after seeding.

6.4.2 Cryoprotection

To determine whether these crystals diffracted, a single crystal was mounted directly into the cryostream of the home source X-ray diffraction equipment without cryoprotection. The crystal diffracted well but ice rings were present in the diffraction pattern, so it was necessary to find a suitable cryoprotectant. It was initially attempted to cryoprotect with glycerol; a crystal was transferred into a solution containing 30% PEG 2K MME, 0.05 M Tris pH 8.0, 0.2 M TMAO and 25% glycerol for 30 sec, and then transferred into the cryostream. The cryoprotection was successful and a complete dataset was collected. The data were processed and a 2mFo-$\triangle$Fc map showed additional density was present in the enzyme active site. The structure was fully refined and glycerol was identified as being bound in the active site. As a small molecule with high structural similarity to the ALDC substrate it was not surprising that glycerol was bound in the active site. However as the aim was to generate structures of ALDC in complex with transition state analogues, glycerol binding made it unsuitable as a cryoprotectant. The concentration of glycerol required to act a cryoprotectant (25%) means it is in vast excess over other ligands, possibly out-competing them for binding in the active site. An alternative cryoprotectant was identified to circumvent this. A crystal was transferred to a drop of LV CryoOil for 30 sec and transferred directly into the cryostream, the crystal was successfully cryoprotected by using the LV CryoOil and this was used as the method of

cryoprotection for all future X-ray diffraction data collection experiments.

6.5 Co-crystallisation with transition state analogues

Separate crystallisation experiments were performed in the presence of all four 2,3-dihydroxy-2-methylbutanoic acids **11**, separately. It was possible to grow crystals in the presence of all four stereoisomers. Data were collected for all four co-crystals but after the data were processed and refined only three of the compounds were successfully bound to the ALDC active site; (2*R*,3*R*)-**11**, (2*S*,3*S*)-**11** and (2*S*,3*R*)-**11**. Crystals grown in the presence of (2*R*,3*S*)-**11** had the enantiomer (2*S*,3*R*)-**11** bound in the active site. Compound (2*S*,3*R*)-**11** was enantiomerically enriched to 75% ee meaning the compound contained 12.5% of the minor enantiomer and this was a sufficient amount to allow it to bind. The binding agrees with the kinetic results presented in Chapter 5. The three competitive inhibitors were able to bind within the active site during the crystallisation process whereas the weakest inhibitor ((2*R*,3*S*)-**11**), a mixed inhibitor, did not bind in the ALDC active site during crystallisation; instead it was out-competed by the other less abundant enantiomer.

Five novel ALDC structures have been generated in this study, complexed with phosphate, glycerol, (2*R*,3*R*)-**11**, (2*S*,3*S*)-**11** and (2*S*,3*R*)-**11**. The crystallographic data collection and refinement statistics for all five structures are contained in Table 6.1. There was little variation in the global fold of all five structures. This was quantified by calculating the global RMSD between each of the structures; the RMSD values are included in Table 6.2. The five structures solved in this project were also compared to the previously solved structures. All six *B. brevis* structures are highly similar with an RMSD values ranging from 0.058-0.135 Å which shows that their are no global changes to the protein structure upon ligand binding. When comparing the *B. brevis* structures to the *S. aureus* structure the RMSD is slightly larger

Table 6.1: Crystallographic data collection and refinement statistics

	ALDC-PO4	ALDC-RR	ALDC-SR	ALDC-SS	ALDC-GOL
Data collection					
X-ray source, detector	Diamond IO4	Diamond IO4	Diamond IO4	marμX-MAR345	marμX-MAR345
and wavelength (Å)	ADSC Q315, 0.9795	ADSC Q315, 0.9795	ADSC Q315, 0.9795	Cu-Ka = 1.15418	Cu-Ka = 1.15418
Unit cell (a=b, c (Å))	47.01, 198.97	47.11, 198.91	47.14, 198.66	46.83, 198.28	47.13, 198.89
Resolution (Å)	32-1.1 (1.16-1.1)	35-1.1 (1.16-1.1)	35-1.1 (1.16-1.1)	41-1.6 (1.69-1.6)	41-1.6 (1.69-1.6)
Observations	512716	524802	711104	689609	484164
Unique reflections	104587	105260	104653	34150	34655
$I/\sigma(I)$	11.0 (3.3)	10.8 (2.4)	11.0 (1.9)	19.5 (4.6)	14.5 (3.4)
R_{sym}[a]	0.069 (0.345)	0.128 (0.761)	0.098 (0.797)	0.193 (0.837)	0.183 (1.010)
Completeness (%)	99.6 (99.8)	99.8 (99.4)	99.3 (95.1)	99.3 (97.4)	99.1 (94.1)
Refinement					
All non-hydrogen atoms	2307	2273	2300	2229	2243
Water molecules	422	384	411	340	331
Other solvent molecules	Phosphate	(2*R*,3*R*)-**11**	(2*S*,3*R*)-**11**	(2*S*,3*S*)-**11**	1+2 Glycerols
R_{cryst}[b]	0.164 (0.204)	0.207 (0.420)	0.177 (0.283)	0.180 (0.258)	0.171 (0.379)
Reflections used	100421 (7161)	101064 (7273)	100480 (6747)	32787 (2281)	33269 (2140)
R_{free}[c]	0.180 (0.219)	0.225 (0.406)	0.184 (0.297)	0.210 (0.258)	0.212 (0.413)
Reflections used	4166 (268)	4196 (272)	4173 (253)	1363 (92)	1386 (90)
R_{cryst} (all data)[b]	0.165	0.207	0.177	0.181	0.172
Mean temperature factor (Å^2)	12.9	10.2	12.5	10.4	12.3
Rmsds from ideal values					
Bonds (Å)	0.015	0.015	0.016	0.016	0.016
Angles (°)	1.6	1.6	1.6	1.5	1.5
DPI coordinate error[d](Å)	0.029	0.036	0.031	0.089	0.084
PBD ID	4BT7	4BT3	4BT5	4BT4	4BT6

Numbers in parentheses refer to values in the highest resolution shell.

[a] $R_{sym} = \Sigma_j \Sigma_h |I_{h,j} - < I_h >| / \Sigma_j \Sigma_h < I_h >$ where $I_{h,j}$ is the jth observation of reflection h and $< I_h >$ is the mean intensity of that reflection.

[b] $R_{cryst} = \Sigma ||F_{obs}| - |F_{calc}|| / \Sigma |F_{obs}|$ where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively.

[c] R_{free} is equivalent to R_{cryst} for a randomly selected subset of reflections not used in the refinement [23].

[d] DPI refers to the diffraction component precision index [124].

Table 6.2: Comparison of the RMSD values between ALDC structures

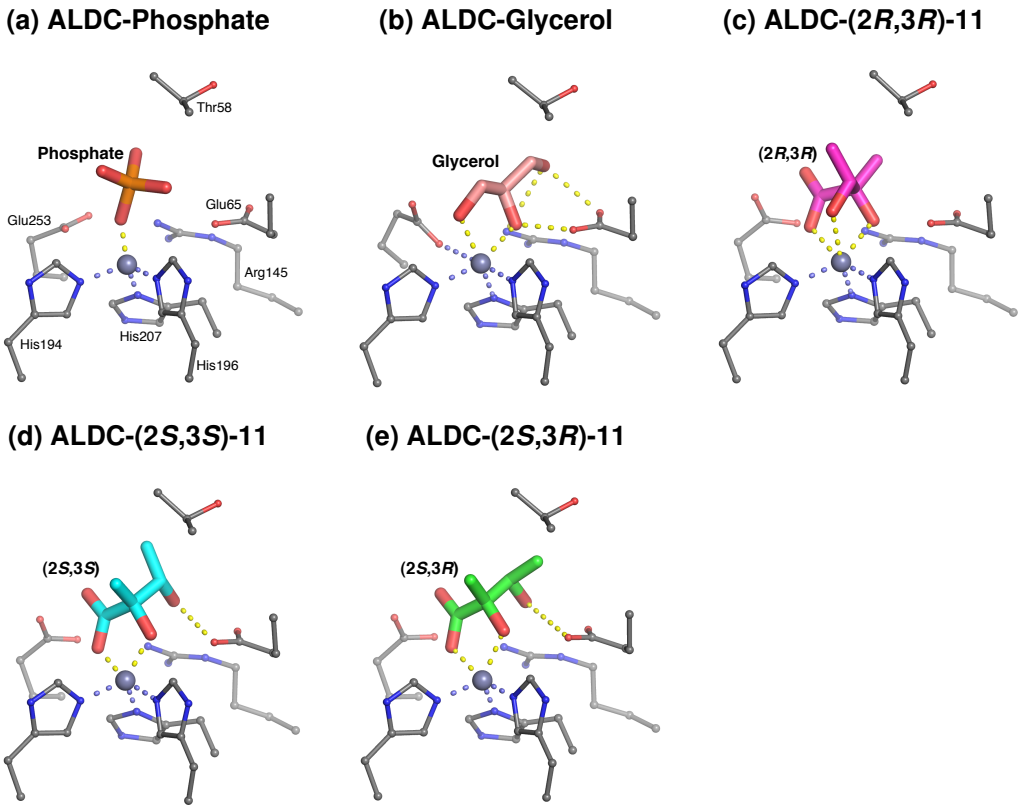
	PO4	GOL	RR	SS	SR	EG	SA
PO4	-	-	-	-	-	-	-
GOL	0.115	-	-	-	-	-	-
RR	0.069	0.107	-	-	-	-	-
SS	0.111	0.135	0.079	-	-	-	-
SR	0.077	0.118	0.058	0.082	-	-	-
EG	0.087	0.093	0.086	0.132	0.109	-	-
SA	0.689	0.690	0.696	0.702	0.691	0.704	-

*The RMSD values are given in angströms (Å). PO4, GOL, RR, SS, SR and EG are the B. brevis structures complexed with phosphate, glycerol, (2R,3R)-**11**, (2S,3S)-**11**, (2S,3R)-**21** and ethylene glycol, respectively. SA is the S. aureus structure.*

at approximately 0.7 Å between all structures but this is still a very minor difference in the global structure. The active site structures of ALDC with phosphate, glycerol, (2*R*,3*R*)-**22**, (2*S*,3*S*)-**11** and (2*S*,3*R*)-**22** are shown in Figure 6.3.

Within these five structures a high degree of flexibility can be observed in the C-terminus. Glu253 can be observed in three different positions; when glycerol is bound (Figure 6.3b) it is coordinated to Zn^{2+}; in the phosphate-bound structure (Figure 6.3a) Glu253 is slightly further from the metal; in all three 2,3-dihydroxy-2-methylbutanoic acid **11** structures (Figure 6.3c, 6.3d and 6.3e) Glu253 is shifted even further from the metal. It is clear from these structures that a shift in the C-terminus is necessary for substrate binding. When ALDC was crystallised in the absence of any ligands, phosphate was observed binding in the ALDC active site as shown in Figure 6.3a. Phosphate was not present in the crystallisation conditions or the buffer the enzyme was stored in; however, phosphate was present in the media during the protein expression. It is likely that phosphate became bound to the enzyme during the expression and remained bound during the purification and storage. Glycerol binds to ALDC through two O-Zn coordination bonds, the third hydroxyl group is within hydrogen bonding

Figure 6.3: Structures of the ALDC active site



*ALDC structures with transition state analogues bound in the active site (a) in the resting state phosphate, His194, 196 and 207 are coordinated to Zn^{2+}. (b) Glycerol (salmon) binds to the ALDC active through two O-Zn coordination bonds, the third OH is interacting with two residues Glu65 and Arg145, in this structure Glu253 is also coordinated to Zn^{2+}. (c) (2R,3R)-2,3-dihydroxy-2-methylbutanoic acid (2R,3R)-**11** (magenta) an (S)-acetolactate (S)-**3** analogue binds through three O-Zn bonds and causes Glu253 to shift position. (d) (2S,3S)-2,3-dihydroxy-2-methylbutanoic acid (2S,3S)-**11** (cyan) is an (R)-acetolactate (R)-**3** analogue and binds through two oxygens, the other OH is pointing away from the zinc and interacts with Glu65. (e) (2S,3R)-2,3-dihydroxy-2-methylbutanoic acid (2S,3R)-**11** (green) is another (R)-acetolactate (R)-**3** analogue which binds almost identically to (2S,3S)-**11** (cyan). For clarity the electron density maps have been omitted however this figure is reproduced in Appendix G with the electron density shown.*

distance to two active site residues Arg145 and Glu65. The enantiomerically enriched transition state analogues bind differently, the (*R*)-acetolactate (*R*)-**3** analogues (2*S*,3*S*)-**11** and (2*S*,3*R*)-**11** Figure 6.3d and 6.3e, respectively, bind in a similar manner to glycerol. They bind through two O-Zn coordination bonds, one from the negatively charged carboxylate and the tertiary alcohol group. The secondary alcohol group is pointed away from the metal and forms a hydrogen bond with Glu65. Compounds (2*S*,3*S*)-**11** and (2*S*,3*R*)-**11** bind almost identically apart from the direction of the methyl group due to the differing chirality at this position. The (*S*)-acetolactate (*S*)-**3** analogue (2*R*,3*R*)-**3** binds differently to the other compounds. Compound (2*R*,3*R*)-**3** binds through three O-Zn coordination bonds through the negatively charged carboxylate and both alcohol groups. The other (*S*)-acetolactate (*S*)-**3** analogue, (2*R*,3*S*)-**11**, was the weakest inhibitor and did not form an ALDC complex. This interaction is unlikely because for (2*R*,3*S*)-**11** to bind in a similar manner to (2*R*,3*R*)-**11** via three O-Zn coordination bonds it would force the methyl group into an unfavourable position, near the backbone carbonyl group of Gly57.

6.6 Proposed mechanism of ALDC

The structures of ALDC complexed with (*S*)-**3** and (*R*)-**3** transition state analogues led to the proposal of the enzyme mechanism shown in Figure 6.4. It is already known from NMR studies [56,58,59] that (*S*)-[3,4-$^{13}C_2$]-acetolactate (*S*)-[3,4-$^{13}C_2$]-**3** is decarboxylated to give (*R*)-[1,2-$^{13}C_2$]-acetoin (*R*)-[1,2-$^{13}C_2$]-**4** with the ^{13}C labels retained in the same positions in the product but that (*R*)-[3,4-$^{13}C_2$]-acetolactate (*R*)-[3,4-$^{13}C_2$]-**3** is decarboxylated to give (*R*)-[3,4-$^{13}C_2$]-acetoin (*S*)-[3,4-$^{13}C_2$]-**4** with the ^{13}C labels switched in the product. Using a combination of the ^{13}C labelling experiments and the structural information presented in Figure 6.3 it is proposed that ALDC is a bifunctional enzyme with both decarboxylase and

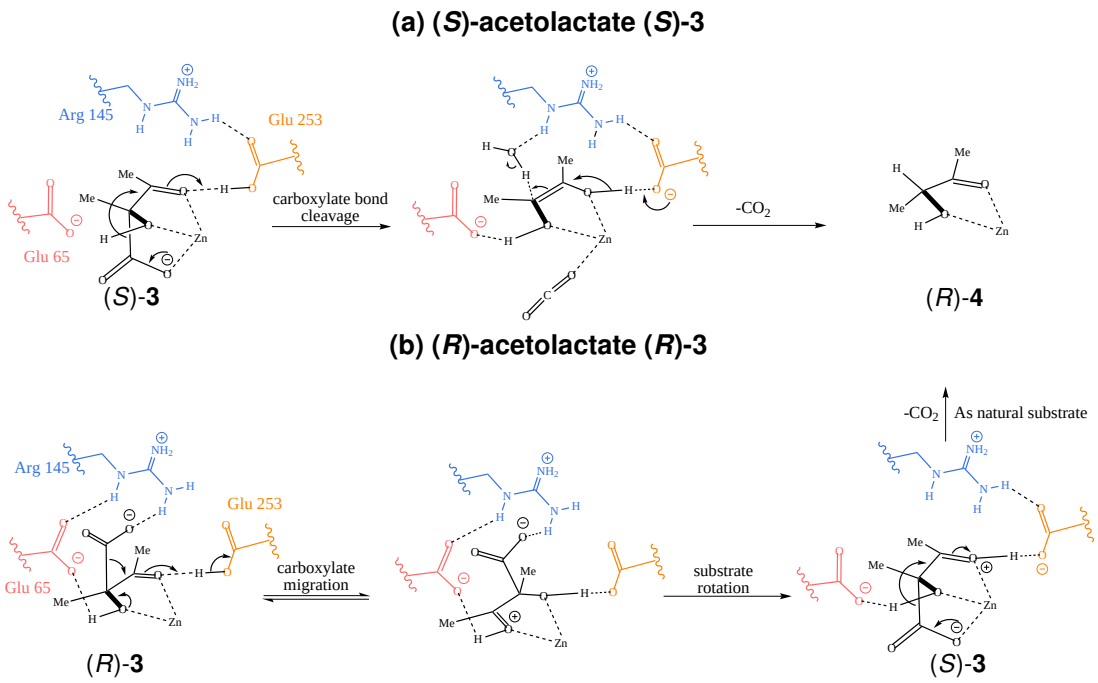
isomerase activity. The natural substrate for ALDC (*S*)-acetolactate (*S*)-**3** binds in the ALDC active site via three O-Zn coordination bonds and it undergoes a Zn^{2+} catalysed decarboxylation to form an enolate intermediate with Glu253 acting as a base. CO_2 is released and the enolate is stereospecifically protonated through assistance of Arg145 as shown in Figure 6.4a. The unnatural substrate (*R*)-acetolactate (*R*)-**3** cannot bind in a conformation to allow decarboxylation to occur; instead it binds via two O-Zn coordination bonds with the carboxylate group pointing away from the metal and interacting with Arg145. Instead of catalysing a decarboxylation the enzyme catalyses an isomerisation, a tertiary ketol rearrangement with carboxylate migration. This rearrangement converts (*R*)-**3** into (*S*)-**3** which can then reorient in the active site as shown in Figure 6.4b and undergo the decarboxylation mechanism shown in Figure 6.4a. Chiral acetolactate **3** has been shown to racemise under basic conditions (pH 12.9 and above) by a tertiary ketol rearrangement with carboxylate migration and it is likely that this rearrangement would also take place within the enzyme active site.

The proposed mechanism agrees with the mutagenesis results presented in Chapter 5. ALDC activity was lost if either Glu65 or Glu253 were mutated to Ala or Gln. Mutation of Arg145 to Ala resulted in a complete loss of activity whereas Lys could retain approximately 40% activity indicating that the enzyme has some tolerance over which base is present in the active site.

6.7 Conclusion

Attempts to produce ALDC crystals under the previously developed conditions had limited success and diffraction quality crystals were grown by optimising a new hit from a fresh commercially available screen. These crystals belong to the same space group as the previous reported Type 2 crystals and were grown under different but similar conditions. The crystals

Figure 6.4: Proposed ALDC mechanism



*It is proposed based on structural information that ALDC is a bifunctional enzyme with decarboxylase and isomerase activity. (a) (S)-acetolactate (S)-**3** is the only substrate for decarboxylation the loss of CO_2 is assisted by Zn^{2+} to form an enolate intermediate. Arg145 (blue) could facilitate protonation of the intermediate to give (R)-acetoin (R)-**4**. (b) (R)-acetolactate (R)-**3** cannot bind in the correct orientation for decarboxylation, instead ALDC catalyses a rearrangement to give (S)-**3** which reorients in the active site to undergo decarboxylation*

used for diffraction studies were grown using the streak seeding method which was found to dramatically improve the quality of the crystals.

It was possible to co-crystallise ALDC in the presence of all four enantiomerically enriched 2,3-dihydroxy-2-methylbutanoic acid **11** and structures of ALDC in complex with (2*S*,3*S*)-**11**, (2*R*,3*R*)-**11** and (2*S*,3*R*)-**11** were obtained at a resolution 1.6 Å, 1.1 Å and 1.1 Å, respectively. It was not possible to get a structure of ALDC in complex with (2*R*,3*S*)-**11** as this was the weakest inhibitor (Chapter 5). The three ALDC-transition state structures led to the proposal of ALDC as a bifunctional enzyme with decarboxylase and isomerase activity. It has been proposed that (*S*)-acetolactate is the only substrate for decarboxylation but that (*R*)-acetolactate cannot bind in a favourable conformation for decarboxylation and instead the enzyme catalyses an isomerisation, which converts (*R*)-acetolactate into (*S*)-acetolactate which can then be decarboxylated.

Chapter 7

General Discussion and Conclusions

ALDC is an unusual enzyme in its ability to convert a racemic mixture into an enantiomerically pure product. Enzymes typically exhibit high stereospecificity but it is rare for an enzyme to show redundancy over the substrate yet have such specificity over the product. A growing field in organic chemistry is asymmetric synthesis and it is still a major challenge to cleanly and efficiently synthesise enantiomeric compounds. As such the use of enzymes as biocatalysts is important in the field of synthetic chemistry. ALDC has the potential to be developed as a biocatalyst for use in the synthesis of enantiomerically pure compounds but this aim has been limited by a lack of understanding of the ALDC mechanism. Therefore the aim of this project was to gain further insight into the mechanism of ALDC. This chapter aims to discuss the work that has been conducted to achieve this aim and put into context the potential future applications.

The proposed mechanism for ALDC was presented in Chapter 6. It was developed through using a combination of techniques: synthetic chemistry, molecular biology, X-ray crystallography and enzyme kinetic studies. The main aim of this project, to gain insight into the mechanism of ALDC, was split into a number of smaller aims.

7.1 Preparation of enantiomerically enriched substrate and transition state analogues

Due to the unusual stereochemistry of ALDC is was necessary to synthesise chiral substrate and transition state analogues prior to the investigation of the ALDC mechanism by kinetic and structural techniques. As both enantiomers of acetolactate are substrates of ALDC that are turned over at different rates, enantiomerically pure substrate was required for kinetic studies. Acetolactate is unstable so must be prepared by hydrolysis of a stable ester precursor. No enantiomerically enriched acetolactate precursors were available commercially and the routes described in the literature involved several steps and were low yielding. Therefore a novel three step synthetic protocol for the production of (*S*)-methyl acetolactate (described in Chapter 3) was developed. (*S*)-Methyl acetolactate was successfully prepared in a 69% yield with an enantiomeric enrichment of >95%. Conditions were also developed to ensure that the hydrolysis of (*S*)-methyl acetolactate gave (*S*)-acetolactate without racemisation occurring.

Also described in Chapter 3 was the design and synthesis of chiral transition state analogues. These were required to give information about the similarities and differences in the interactions of (*S*)- and (*R*)-acetolactate. 2,3-Dihydroxy-2-methylbutanoic acid was identified as a suitable target, containing a hydroxyl group to mimic the departing carboxylate in the transition state and a carboxylate to mimic the ketone with a partial negative charge. As 2,3-dihydroxy-2-methylbutanoic acid contains two chiral centres it exists in four stereoisomeric forms. The four stereoisomers of 2,3-dihydroxy-2-methylbutanoic acid were prepared in a two step synthesis; an asymmetric dihydroxylation followed by ester hydrolysis. The compounds were isolated in reasonable yields with a good enantiomeric enrichment for use

in enzyme inhibition studies and to generate structures of ALDC-transition state analogue complexes.

7.2 Preparation of recombinant AlsD and BdhA

Chapter 4 described the use of molecular biology techniques to clone, over-express and purify recombinant *Bacillus subtilis* AlsD and BdhA. *B. subtilis* BdhA was required for use in the development of a coupled assay for the activity of ALDC, it was successfully cloned into an expression vector with a N-terminal histine tag. BdhA was over-expressed in *E. coli* B834(DE3) cells and purified by IMAC.

Although previous crystallography studies had been carried out on *Bacillus brevis* ALDC and a sample of the native protein was available for use in this project, there was no provision for the generation of active site mutants. Therefore it was decided to engineer an expression vector encoding ALDC that could be used as a template to generate active site mutants. As *B. subtilis* genomic DNA was readily available, it was decided to clone the *B. subtilis alsD* gene into an expression vector with a N-terminal histidine tag. Eight active site point mutants were generated: T55A, T55S, E62A, E62Q, R142A, R142K, E251A and E251Q. Native and mutant AlsD were over-expressed in *E. coli* B834(DE3), the nine protein samples were purified by IMAC however two lower molecular weight fragments were identified by SDS-PAGE. Further chromatographic separation by size exclusion, hydrophobic interaction and ion exchange were unable to resolve the multiple species. The lower molecular weight species were identified by mass spectrometry and western blot to be N-terminally truncated AlsD. The inability to separate the multiple species and the apparent loss of the histidine tag implies that native species contained a mixture of dimers containing truncated AlsD:full length AlsD and full length AlsD:full length AlsD. Although truncated AlsD was present, it

was a minor impurity and the proteins were used in enzyme kinetics studies.

7.3 Kinetic studies of ALDC

Attempts to study the kinetics of ALDC activity using the Voges-Proskauer assay were unsuccessful. The VP assay produced inconsistent results due to unstable substrate solution, unstable colour reagent and a background reaction. Two alternative assays were developed to overcome these challenges. A continuous coupled assay was developed which linked the ALDC catalysed reaction to BdhA activity. As BdhA is a NADH dependent enzyme its activity could be monitored by UV spectroscopy, by following the loss of NADH signal at 340 nm. This assay was suitable for ALDC activity studies and the large extinction coeffient of NADH meant it was sensitive to low substrate concentrations. However attempts to study ALDC inhibition failed as the inhibitors were also inhibiting the coupling enzyme. A second continuous assay was developed to overcome this challenge. The inherent chirality of the ALDC catalysed reaction enabled the use of CD to study ALDC activity. The formation of (*R*)-acetoin was followed at 278 nm, a wavelength where only product was detectable. This assay was less sensitive than the coupled assay, requiring higher concentrations of substrate, but this was outweighed by the quality of the data measured. Activity studies of the active site mutants identified that mutation of E62 or E251 resulted in the complete loss of activity. R142A was also catalytically inactive but R142K retained partial activity and the T55 mutants were more active than native AlsD. These results suggested that the two glutamate residues were essential and a basic residue was necessary for catalysis. ALDC activity in the presence of the chiral transition state analogues identified (2*R*,3*R*)-**11**, (2*S*,3*S*)-**11** and (2*S*,3*R*)-**11** as competitive inhibitors and (2*R*,3*S*)-**11** as a mixed inhibitor.

7.4 Structural studies of ALDC

ALDC was co-crystallised in the presence of three chiral transition state analogues; (2*R*,3*R*)-**11**, (2*S*,3*S*)-**11** and (2*S*,3*R*)-**11**. The structures were solved to 1.1 Å, 1.6 Å and 1.1 Å respectively. In addition structures were obtained with inorganic phosphate and glycerol bound in the active site. These data are the first evidence to describe how the two enantiomers of substrate bind in the active site. The (*S*)-acetolactate analogue ((2*R*,3*R*)-**11**) binds through three O-Zn coordination bonds with the carboxylate coordinated to the metal ion. The (*R*)-acetolactate analogues ((2*S*,3*S*)-**11** and (2*S*,3*R*)-**11**) bind through two O-Zn coordination bonds with the carboxylate pointing away from the metal ion. These data in combination with the active site mutant activity study and the previous ^{13}C labelling studies [54,56,125] allowed the proposal of a mechanism for ALDC activity as described in Chapter 6. These data support the mechanism proposed by Crout et al. [62] that ALDC catalyses an isomerisation with carboxylate migration in addition to a decarboxylation.

7.5 Structure and mechanism of ALDC

ALDC is a metal-dependent decarboxylase. Metal ion decarboxylases can fall into two major categories: oxidative and non-oxidative. ALDC is a non-oxidative Zn^{2+} dependent decarboxylase which puts it into the same category as ACMSD. The active sites between the two enzymes are similar. Both enzymes contain a Zn^{2+} ion coordinated to three histidines and an acidic residue: aspartate in ACMSD and glutamate in ALDC. Both enzymes also contain an arginine residue which in ACMSD forms a H-bond network with active site water molecules. The major difference between the two active sites is that ACMSD has a fourth histidine residue in the active site rather than the additional glutamate residue present in

ALDC. Although ACMSD and ALDC have similar active sites, the mechanisms of the two decarboxylases are very different. In ALDC the active site metal coordinates substrate whereas in ACMSD the metal coordinates a catalytic hydroxide ion [40].

The ALDC mechanism proceeds through the formation of a metal coordinated enediol intermediate which has similarity to the enediol intermediate of the Zn^{2+} class II fructose-1,6-bisphosphate aldolase (FBP-aldolase). The FBP-aldolase active site contains a Zn^{2+} ion coordinated to three histidines with two acidic residues and an arginine nearby [126].

The ALDC mechanism involves a stereoselective protonation of the enolate intermediate and a similar step occurs in the mechanism of malonate decarboxylase. Malonate decarboxylase is cofactor independent, yet it forms an enolate intermediate after decarboxylation which is stereoselectively protonated with inversion of configuration, as in ALDC [127,128].

ALDC is a bifunctional enzyme that catalyses two different reactions within the same active site. In this case the natural reaction is a decarboxylation but the enzyme is also capable of catalysing an isomerisation reaction. Other examples of bifunctional enzymes are isocitrate decarboxylase and ketol-acid reductoisomerase. Isocitrate decarboxylase, a Mn^{2+} dependent enzyme, catalyses the oxidation of isocitrate to oxalosuccinate followed by the decarboxylation to α-keto glutarate [129]. Ketol-acid reductoisomerase also catalyses two reactions; the first is a rearrangement of acetolactate, however in contrast to the mechanism of ALDC, ketol acid reductoisomerase catalyses a methyl migration to form 3-hydroxy-3-methyl-2-ketobutyrate which is subsequently reduced in a $NADP^+$ dependent manner to give 2,3-dihydroxy-3-methylbutyrate. What is unusual about ALDC is the isomerisation reaction does not appear to be biologically important. In the butanediol fermentation pathway the upstream enzyme acetolactate synthase has been shown to produce (*S*)-acetolactate [130] so it is unlikely that ALDC encounters (*R*)-acetolactate naturally. The rearrangement of acetolactate by carboxylate migration under base catalysed conditions is well established

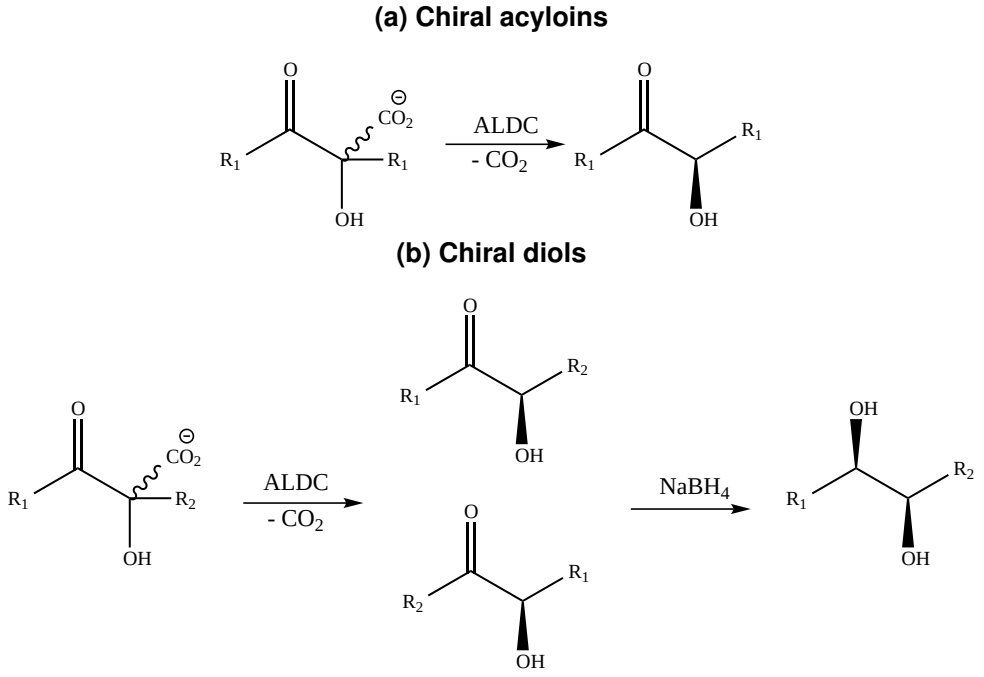
[54,61] and it is more likely that acetolactate is prone to this rearrangement reaction, which can be accommodated within the ALDC active site. It was hypothesised that the two catalytic activities may be distinct and an ALDC mutant that is incapable of decarboxylation may still be able to catalyse isomerisation. This would provide further proof of the mechanism. In this study all of the mutants produced that were catalytically incapable of producing (*R*)-acetoin were also incapable of racemising (*S*)-acetolactate. This was determined by monitoring the activity using the CD assay at two wavelengths, to observe substrate loss and product gain.

7.6 Potential biocatalyst development

The future aim of this work would be to develop ALDC as a biocatalyst. It has already been used in an *E.coli* whole cell biocatalyst for the production of butanediol [53] but there is the potential to use it for the production of enantiomerically enriched chiral acyloins or diols, as shown in Figure 7.1.

Asymmetric biocatalysts are used in asymmetric synthesis to yield a product with 100% enantiomeric enrichment. There are two main methods; kinetic resolution and enantioselective desymmetrisation. Kinetic resolution takes advantage of the preference of an enzyme for one enantiomer of substrate; typically one enantiomer of substrate is turned over to give a single enantiomer of product, the remaining enantiomer of substrate can then be separated from the product. Kinetic resolution is limited to a 50% maximal yield but if linked to a racemisation reaction of the substrate, as in dynamic kinetic resolution, the yield can be improved. Enantioselective desymmetrisation utilises an enzyme that can convert a prochiral or symmetric substrate (meso) into a chiral product [131]. The reaction catalysed by ALDC is essentially a dynamic kinetic resolution with both steps catalysed by a single catalyst. Some preliminary studies into the substrate range of ALDC have been conducted. The natural

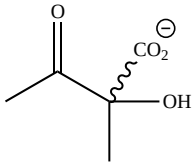
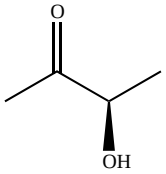
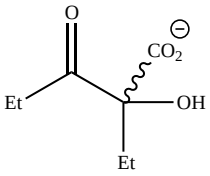
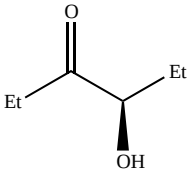
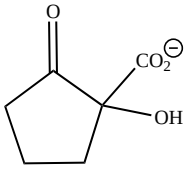
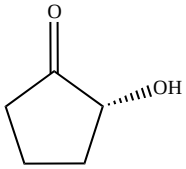
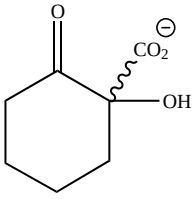
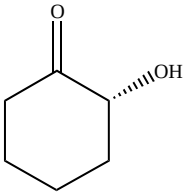
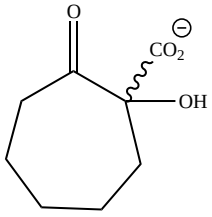
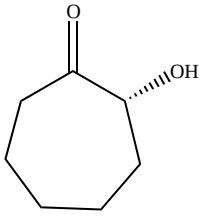
Figure 7.1: Application of ALDC for the production of chiral acyloins or diols



ALDC has the potential to be developed as a biocatalyst for the production of enantiomerically enriched compounds. (a) If a substrate with R_1=R_1 is used, then a single enantiomeric acyloin will be produced. (b) If a substrate with $R_1 \neq R_2$ is used then a diastereomeric mixture of acyloins will be produced. A subsequent diastereoselective reduction would give an enantiomerically enriched diol.

substrates of ALDC are (*S*)-acetolactate and (*S*)-acetohydroxybutyrate, but ALDC appears to have a much broader substrate range than those naturally encountered. Table 7.1 shows the non-natural ALDC substrates that have been identified by Crout et al. [60]. This broad substrate specificity and the increased activity of the *B. subtilis* T55S and T55A mutants indicates it may be possible to further extend the substrate scope of ALDC by site-directed mutagenesis of the active site. These kinds of experiments require the use of a robust assay to test the activity with different substrates. This may not be possible with the coupled assay as it requires the ALDC product to be turned over by the coupling enzyme whereas the CD assay would be suitable as it could detect the chirality of different products.

Table 7.1: ALDC substrate range

Substrate	Product	Yield (%)	Ee (%)
 **3**	 **4**	70	> 96
 **23**	 **24**	90	83
 **25**	 **26**	90	82
 **27**	 **28**	67	79
 **29**	 **30**	90	83

ALDC is able to decarboxylate a broader range of substrates than those naturally occurring. This table has been adapted from Crout et al. [60].

7.7 Conclusion

ALDC is a unique enzyme that converts racemic acetolactate into enantiomerically pure (*R*)-acetoin. The work presented in this thesis has provided further evidence which indicates that ALDC is a bifunctional enzyme with isomerase and decarboxylase activity. By combining the mutant activity studies with structures of ALDC in complex with transition state analogues it has been possible to propose a molecular mechanism to explain the dual functionality of ALDC. ALDC decarboxylates the favoured substrate, (*S*)-acetolactate, to give enantiomerically pure (*R*)-acetoin whereas the less favoured substrate, (*R*)-acetolactate undergoes an isomerisation with carboxylate migration followed by decarboxylation [64].

References

[1] L. Pauling, 1948. “Chemical achievement and hope for the future”, *Am. Sci.* Vol. 36, pages 51–58.

[2] V. L. Schramm, 1998. “Enzymatic transition states and transition state analog design”, *Annual Review of Biochemistry*, vol. 67, no. 1, pages 693–720.

[3] R. Wolfenden, 1999. “Conformational aspects of inhibitor design: enzyme-substrate interactions in the transition state”, *Bioorganic & Medicinal Chemistry*, vol. 7, no. 5, pages 647–652.

[4] V. L. Schramm, 2005. “Enzymatic transition states and transition state analogues”, *Current Opinion in Structural Biology*, vol. 15, no. 6, pages 604–613.

[5] J. J. Ruiz-Pernía, E. Silla, and I. Tuñón, 2007. “Enzymatic effects on reactant and transition states. The case of chalcone isomerase”, *Journal of the American Chemical Society*, vol. 129, no. 29, pages 9117–9124.

[6] V. L. Schramm, 2011. “Enzymatic transition states, transition-state analogs, dynamics, thermodynamics, and lifetimes”, *Annual Review of Biochemistry*, vol. 80, pages 703–732.

[7] Z. S. Olempska-Beer, R. I. Merker, M. D. Ditto, and M. J. DiNovi, 2006. “Food-processing enzymes from recombinant microorganismsa review”, *Regulatory Toxicology and Pharmacology*, vol. 45, no. 2, pages 144–158.

[8] C. W. Bamforth, 2009. “Current perspectives on the role of enzymes in brewing”, *Journal of Cereal Science*, vol. 50, no. 3, pages 353–357.

[9] K. R. Jegannathan, E.-S. Chan, and P. Ravindra, 2011. “Biotechnology in Biofuels-A Cleaner Technology”, *Journal of Applied Sciences*, vol. 11, no. 13, pages 2421–2425.

[10] F. Hasan, A. A. Shah, S. Javed, and A. Hameed, 2010. “Enzymes used in detergents: Lipases”, *Afr J Biotechnol*, vol. 9, pages 4836–4844.

[11] G. A. Behrens, A. Hummel, S. K. Padhi, S. Schätzle, and U. T. Bornscheuer, 2011. “Discovery and protein engineering of biocatalysts for organic synthesis”, *Advanced Synthesis & Catalysis*, vol. 353, no. 13, pages 2191–2215.

[12] L. Michaelis and M. L. Menten, 1913. “Die kinetik der invertinwirkung”, *Biochemische Zeitschrift*, vol. 49, no. 352, pages 333–369.

[13] K. A. Johnson and R. S. Goody, 2011. “The original Michaelis constant: translation of the 1913 Michaelis–Menten paper”, *Biochemistry*, vol. 50, no. 39, pages 8264–8269.

[14] R. A. Alberty and G. G. Hammes, 1958. “Application of the theory of diffusion-controlled reactions to enzyme kinetics”, *The Journal of Physical Chemistry*, vol. 62, no. 2, pages 154–159.

[15] H. Lineweaver and D. Burk, 1934. “The determination of enzyme dissociation constants”, *Journal of the American Chemical Society*, vol. 56, no. 3, pages 658–666.

[16] A. A. Saboury, 2009. “Enzyme inhibition and activation: a general theory”, *Journal of the Iranian Chemical Society*, vol. 6, no. 2, pages 219–229.

[17] C Haas and J Drenth, 1999. “Understanding protein crystallization on the basis of the phase diagram”, *Journal of Crystal Growth*, vol. 196, no. 2, pages 388–394.

[18] N. Asherie, 2004. “Protein crystallization and phase diagrams”, *Methods*, vol. 34, no. 3, pages 266–272.

[19] N. E. Chayen and E. Saridakis, 2008. “Protein crystallization: from purified protein to diffraction-quality crystal”, *Nature Methods*, vol. 5, no. 2, pages 147–153.

[20] T. G. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell, and A. G. Leslie, 2011. “iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM”, *Acta Crystallographica Section D: Biological Crystallography*, vol. 67, no. 4, pages 271–281.

[21] Z. Dauter, 1999. “Data-collection strategies”, *Acta Crystallographica Section D: Biological Crystallography*, vol. 55, no. 10, pages 1703–1717.

[22] H. Hauptman, 1997. “Phasing methods for protein crystallography”, *Current Opinion in Structural Biology*, vol. 7, no. 5, pages 672–680.

[23] A. T. Brünger, 1992. “Free R value: a novel statistical quantity for assessing the accuracy of crystal structures”, *Nature*, vol. 355, pages 472–475.

[24] T. Li, L. Huo, C. Pulley, and A. Liu, 2012. “Decarboxylation mechanisms in biological system”, *Bioorganic Chemistry*, vol. 43, pages 2–14.

[25] F. H. Westheimer, 1995. “Coincidences, decarboxylation, and electrostatic effects”, *Tetrahedron*, vol. 51, no. 1, pages 3–20.

[26] M. C. Ho, J. F. Ménétret, H. Tsuruta, and K. N. Allen, 2009. “The origin of the electrostatic perturbation in acetoacetate decarboxylase”, *Nature*, vol. 459, no. 7245, pages 393–397.

[27] A. Radzicka and R. Wolfenden, 1995. “A proficient enzyme”, *Science*, vol. 267, no. 5194, pages 90–93.

[28] K. K. Chan, B. M. K. Wood, A. A. Fedorov, E. V. Fedorov, H. J. Imker, T. L. Amyes, J. P. Richard, S. C. Almo, and J. A. Gerlt, 2009. “Mechanism of the orotidine

5-monophosphate decarboxylase-catalyzed reaction: evidence for substrate destabilization", *Biochemistry*, vol. 48, no. 24, pages 5518–5531.

[29] M. Fujihashi, K. Mito, E. F. Pai, and K. Miki, 2013. "Atomic-resolution structure of the orotidine 5'-monophosphate decarboxylase product complex combined with surface plasmon resonance analysis: Implications for the catalytic mechanism", *Journal of Biological Chemistry*, vol. 288, pages 9011–9016.

[30] P Arjunan, T Umland, F Dyda, S Swaminathan, W Furey, M Sax, B Farrenkopf, Y Gao, D Zhang, and F Jordan, 1996. "Crystal structure of the thiamin diphosphate-dependent enzyme pyruvate decarboxylase from the yeast *Saccharomyces cerevisiae* at 2.3 Å resolution", *Journal of Molecular Biology*, vol. 256, no. 3, pages 590–600.

[31] H. B. Brooks and M. A. Phillips, 1997. "Characterization of the reaction mechanism for *Trypanosoma brucei* ornithine decarboxylase by multiwavelength stopped-flow spectroscopy", *Biochemistry*, vol. 36, no. 49, pages 15147–15155.

[32] H. Hayashi, F. Tsukiyama, S. Ishii, H. Mizuguchi, and H. Kagamiyama, 1999. "Acid-base chemistry of the reaction of aromatic L-amino acid decarboxylase and dopa analyzed by transient and steady-state kinetics: Preferential binding of the substrate with its amino group unprotonated", *Biochemistry*, vol. 38, no. 47, pages 15615–15622.

[33] D. E. Graham, H. Xu, and R. H. White, 2002. "*Methanococcus jannaschii* uses a pyruvoyl-dependent arginine decarboxylase in polyamine biosynthesis", *Journal of Biological Chemistry*, vol. 277, no. 26, pages 23500–23507.

[34] W. D. Tolbert, J. L. Ekstrom, I. I. Mathews, J. A. Secrist, P. Kapoor, A. E. Pegg, and S. E. Ealick, 2001. "The structural basis for substrate specificity and inhibition of human *S*-adenosylmethionine decarboxylase", *Biochemistry*, vol. 40, no. 32, pages 9484–9494.

[35] T. A. Alston and R. H. Abeles, 1987. "Reaction of *Lactobacillus* histidine decarboxylase with L-histidine methyl ester", *Biochemistry*, vol. 26, no. 13, pages 4082–4085.

[36] B. T. Saylor, L. A. Reinhardt, Z. Lu, M. S. Shukla, L. Nguyen, W. W. Cleland, A. Angerhofer, K. N. Allen, and N. G. J. Richards, 2012. "A structural element that facilitates proton-coupled electron transfer in oxalate decarboxylase", *Biochemistry*, vol. 51, no. 13, pages 2911–2920.

[37] W. S. Yew, J. Akana, E. L. Wise, I. Rayment, and J. A. Gerlt, 2005. "Evolution of enzymatic activities in the orotidine 5'-monophosphate decarboxylase suprafamily: enhancing the promiscuous D-arabino-hex-3-ulose 6-phosphate synthase reaction catalyzed by 3-keto-L-gulonate 6-phosphate decarboxylase", *Biochemistry*, vol. 44, no. 6, pages 1807–1815.

[38] G. L. Waldrop, B. F. Braxton, J. L. Urbauer, W. W. Cleland, and D. M. Kiick, 1994. "Secondary ^{18}O and primary ^{13}C isotope effects as a probe of transition-state structure for enzymic decarboxylation of oxalacetate", *Biochemistry*, vol. 33, no. 17, pages 5262–5267.

[39] A. Liu and H. Zhang, 2006. "Transition metal-catalyzed nonoxidative decarboxylation reactions", *Biochemistry*, vol. 45, no. 35, pages 10407–10411.

[40] L. Huo, A. J. Fielding, Y. Chen, T. Li, H. Iwaki, J. P. Hosler, L. Chen, Y. Hasegawa, L. Que Jr, and A. Liu, 2012. “Evidence for a dual role of an active site histidine in α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase”, *Biochemistry*, vol. 51, no. 29, pages 5811–5821.

[41] J. P. Løken and F. C. Størmer, 1970. “Acetolactate decarboxylase from *Aerobacter aerogenes*”, *European Journal of Biochemistry*, vol. 14, no. 1, pages 133–137.

[42] S. E. Godtfredsen, H. Lorck, and P. Sigsgaard, 1983. “On the occurrence of α-acetolactate decarboxylases among microorganisms”, *Carlsberg Research Communications*, vol. 48, no. 3, pages 239–247.

[43] L. Johansen, K. Bryn, and F. C. Størmer, 1975. “Physiological and biochemical role of the butanediol pathway in *Aerobacter (Enterobacter) aerogenes*.”, *Journal of Bacteriology*, vol. 123, no. 3, pages 1124–1130.

[44] K. Blomqvist, M Nikkola, P Lehtovaara, M. L. Suihko, U Airaksinen, K. B. Stråby, J. K. Knowles, and M. E. Penttilä, 1993. “Characterization of the genes of the 2,3-butanediol operons from *Klebsiella terrigena* and *Enterobacter aerogenes*.”, *Journal of Bacteriology*, vol. 175, no. 5, pages 1392–1404.

[45] N. Goupil-Feuillerat, M. Cocaign-Bousquet, J. Godon, S. D. Ehrlich, and P. Renault, 1997. “Dual role of alpha-acetolactate decarboxylase in *Lactococcus lactis* subsp. *lactis*.”, *Journal of Bacteriology*, vol. 179, no. 20, pages 6285–6293.

[46] B. Diderichsen, U. Wedsted, L. Hedegaard, B. R. Jensen, and C. Sjøholm, 1990. “Cloning of *aldB*, which encodes alpha-acetolactate decarboxylase, an exoenzyme from *Bacillus brevis*.”, *Journal of Bacteriology*, vol. 172, no. 8, pages 4315–4321.

[47] C. Dulieu and D. Poncelet, 1999. “Spectrophotometric assay of α-acetolactate decarboxylase”, *Enzyme and Microbial Technology*, vol. 25, no. 6, pages 537–542.

[48] S. E. Godtfredsen and M. Ottesen, 1982. “Maturation of beer with α-acetolactate decarboxylase”, *Carlsberg Research Communications*, vol. 47, no. 2, pages 93–102.

[49] S. E. Godtfredsen, A. M. Rasmussen, M. Ottesen, P. Rafn, and N. Peitersen, 1984. “Occurrence of α-acetolactate decarboxylases among lactic acid bacteria and their utilization for maturation of beer”, *Applied Microbiology and Biotechnology*, vol. 20, no. 1, pages 23–28.

[50] S. E. Godtfredsen, A. M. Rasmussen, M Ottesen, T Mathiasen, and B Ahrenst-Larsen, 1984. “Application of the acetolactate decarboxylase from *Lactobacillus casei* for accelerated maturation of beer”, *Carlsberg Research Communications*, vol. 49, no. 1, pages 69–74.

[51] C. Dulieu, M. Moll, J. Boudrant, and D. Poncelet, 2000. “Biocatalysts and bioreactor design-improved performances and control of beer fermentation using encapsulated α-acetolactate decarboxylase and modeling”, *Biotechnology Progress*, vol. 16, no. 6, pages 958–965.

[52] K. Blomqvist, M.-L. Suihko, J. Knowles, and M. Penttilä, 1991. “Chromosomal integration and expression of two bacterial α-acetolactate decarboxylase genes in brewer’s yeast”, *Applied and Environmental Microbiology*, vol. 57, no. 10, pages 2796–2803.

[53] Y. Yan, C. C. Lee, and J. C. Liao, 2009. “Enantioselective synthesis of pure *(R,R)*-2,3-butanediol in *Escherichia coli* with stereospecific secondary alcohol dehydrogenases”, *Organic & Biomolecular Chemistry*, vol. 7, no. 19, pages 3914–3917.

[54] D. H. G. Crout and C. J. R. Hedgecock, 1979. “The base-catalysed rearrangement of α-acetolactate (2-hydroxy-2-methyl-3-oxobutanoate): a novel carboxylate ion migration in a tertiary ketol rearrangement”, *Journal of the Chemical Society, Perkin Transactions 1*, no. 0, pages 1982–1989.

[55] R. K. Hill, S. Sawada, and S. M. Arfin, 1979. “Stereochemistry of valine and isoleucine biosynthesis IV. Synthesis, configuration, and enzymatic specificity of α-acetolactate and α-aceto-α-hydroxybutyrate”, *Bioorganic Chemistry*, vol. 8, no. 2, pages 175–189.

[56] D. H. G. Crout, J. Littlechild, M. B. Mitchell, and S. M. Morrey, 1984. “Stereochemistry of the decarboxylation of α-acetolactate (2-hydroxy-2-methyl-3-oxobutanoate) by the acetolactate decarboxylase of *Klebsiella aerogenes*”, *Journal of the Chemical Society, Perkin Transactions 1*, no. 0, pages 2271–2276.

[57] A. F. Drake, G. Siligardi, D. H. G. Crout, and D. L. Rathbone, 1987. “Applications of vibrational infrared circular dichroism to biological problems: stereochemistry of proton exchange in acetoin (3-hydroxybutan-2-one) catalysed by acetolactate decarboxylase”, *Journal of the Chemical Society, Chemical Communications*, no. 24, pages 1834–1835.

[58] D. H. G. Crout and D. L. Rathbone, 1988. “Biotransformations with acetolactate decarboxylase: unusual conversions of both substrate enantiomers into products of high optical purity”, *Journal of the Chemical Society, Chemical Communications*, no. 2, pages 98–99.

[59] D. H. G. Crout, E. R. Lee, and D. L. Rathbone, 1990. “Absolute configuration of the product of the acetolactate synthase reaction by a novel method of analysis using acetolactate decarboxylase”, *Journal of the Chemical Society, Perkin Transactions 1*, no. 5, pages 1367–1369.

[60] D. H. G. Crout, S Davies, R. J. Heath, C. O. Miles, D. R. Rathbone, B. E. P. Swoboda, and M. B. Gravestock, 1994. “Applications of hydrolytic and decarboxylating enzymes in biotransformations”, *Biocatalysis and Biotransformation*, vol. 9, no. 1-4, pages 1–30.

[61] F. B. Armstrong, C. J. R. Hedgecock, J. B. Reary, D. Whitehouse, and D. H. G. Crout, 1974. “Stereochemistry of the reductoisomerase and $\alpha\beta$-dihydroxyacid dehydratase-catalysed steps in valine and isoleucine biosynthesis. Observation of a novel tertiary ketol rearrangement”, *Journal of the Chemical Society, Chemical Communications*, no. 9, pages 351–352.

[62] D. H. G. Crout, J. Littlechild, and S. M. Morrey, 1986. “Acetoin metabolism: stereochemistry of the acetoin produced by the pyruvate decarboxylase of wheat germ and by the

α-acetolactate decarboxylase of *Klebsiella aerogenes*", *Journal of the Chemical Society, Perkin Transactions 1*, pages 105–108.

[63] S. Najmudin, J. T. Andersen, S. A. Patkar, T. V. Borchert, D. H. G. Crout, and V. Fülöp, 2003. "Purification, crystallization and preliminary X-ray crystallographic studies on acetolactate decarboxylase", *Acta Crystallographica Section D: Biological Crystallography*, vol. 59, no. 6, pages 1073–1075.

[64] V. A. Marlow, D. Rea, S. Najmudin, M. Wills, and V. Fu lo p, 2013. "Structure and mechanism of acetolactate decarboxylase", *ACS Chemical Biology*, vol. 8, no. 10, pages 2339–2344.

[65] K. Stritzke, S. Schulz, and R. Nishida, 2002. "Absolute configuration and synthesis of β-and δ-lactones present in the pheromone system of the giant white butterfly *Idea leuconoe*", *European Journal of Organic Chemistry*, vol. 2002, no. 22, pages 3884–3892.

[66] K. Kawai, R. Nishida, and H. Fukami, 1999. "Clerodendrin I, a new neoclerodane diterpenoid from *Clerodendron trichotomum*.", *Bioscience, Biotechnology, and Biochemistry*, vol. 63, no. 10, pages 1795–1797.

[67] A. Battaglia, E. Baldelli, G. Barbaro, P. Giorgianni, A. Guerrini, M. Monari, and S. Selva, 2002. "Synthesis and desilylation of (2*R*,3*S*)-α-methyl-α-silyl-α,β-2,3-dihydroxycarboxylic methyl esters", *Tetrahedron: Asymmetry*, vol. 13, no. 16, pages 1825–1832.

[68] W. Ladner, 1983. "Stereoselektive synthese von alkoholen, XV. Stereoselektive synthese von α-Alkyl-α,β-dihydroxy-carbonsäureester-derivaten: synthese der (+)-viridiflorsäure", *Chemische Berichte*, vol. 116, no. 10, pages 3413–3426.

[69] A. R. Moen, K. Ruud, and T. Anthonsen, 2008. "Combination of stereospecific dihydroxylation and enzyme catalyzed enantioselective resolution for synthesis of enantiopure vicinal diols", *Journal of Molecular Catalysis B: Enzymatic*, vol. 50, no. 2-4, pages 74–79.

[70] B. W. Christensen and A. Kjaer, 1962. "The absolute configuration of 2-hydroxy-2-methylbutyric acid", *Acta Chemica Scandinavica*, vol. 16, pages 2466–2467.

[71] A. Battaglia, G. Barbaro, P. Giorgianni, A. Guerrini, C. Bertucci, and S. Geremia, 2000. "Addition reactions of aldehydes to lithium enolates of 1,3-dioxolan-4-ones: a configurational reassessment", *Chemistry-A European Journal*, vol. 6, no. 19, pages 3551–3557.

[72] D. J. Robins and D. H. G. Crout, 1970. "Pyrrolizidine alkaloids. The absolute configuration at C-2 in monocrotalic acid", *Journal of the Chemical Society C: Organic*, no. 9, pages 1334–1336.

[73] T. Kaneko, M. Sakamoto, K. Ohtani, A. Ito, R. Kasai, K. Yamasaki, and W. G. Padorina, 1995. "Secoiridoid and flavonoid glycosides from *Gonocaryum calleryanum*", *Phytochemistry*, vol. 39, no. 1, pages 115–120.

[74] B. W. Christensen and A. Kjaer, September 1962. "The absolute configuration of *threo*-$\alpha\beta$-dihydroxy-α-methylbutyric acid", *Proceedings of the Chemical Society*, pages 307–308.

[75] Y. Y. Chan, Y. L. Leu, F. W. Lin, C. Y. Li, Y. C. Wu, L. S. Shi, M. J. Liou, and T. S. Wu, 1998. "A secoiridoid and other constituents of *Gonocaryum calleryanum*", *Phytochemistry*, vol. 47, no. 6, pages 1073–1077.

[76] A. Greiner and J. Y. Ortholand, 1992. "Erythroselective aldol condensation of amine free 2-*t*-butyl-5-methyl-2-phenyl-1, 3-dioxolan-4-one lithium enolate synthesis of the ethyl acetolactate enantiomers", *Tetrahedron Letters*, vol. 33, no. 14, pages 1897–1900.

[77] R. J. Abraham, J. J. Byrne, L. Griffiths, and M. Perez, 2006. "^{1}H chemical shifts in NMR: Part 23, the effect of dimethyl sulphoxide versus chloroform solvent on ^{1}H chemical shifts", *Magnetic Resonance in Chemistry*, vol. 44, no. 5, pages 491–509, ISSN: 1097-458X.

[78] D. H. G. Crout and D. L. Rathbone, 1989. "Synthesis of 2-hydroxy-3-oxocarboxylic esters from the corresponding α,β-unsaturated esters by a simple one-step procedure", *Synthesis*, vol. 1989, no. 01, pages 40–42.

[79] F. B. Armstrong, E. L. Lipscomb, D. H. G. Crout, M. B. Mitchell, and S. R. Prakash, 1983. "Biosynthesis of valine and isoleucine: synthesis and biological activity of (2*S*)-α-acetolactic acid (2-hydroxy-2-methyl-3-oxobutanoic acid), and (2*R*)-and (2*S*)-α-acetohydroxybutyric acid (2-ethyl-2-hydroxy-3-oxobutanoic acid)", *Journal of the Chemical Society, Perkin Transactions 1*, pages 1197–1201.

[80] J. Christoffers, T. Kauf, T. Werner, and M. Rössle, 2006. "Cerium-catalyzed α-hydroxylation reactions of α-cyclopropyl β-dicarbonyl compounds with molecular oxygen", *European Journal of Organic Chemistry*, vol. 2006, no. 11, pages 2601–2608.

[81] U. K. Laemmli, 1970. "Cleavage of structural proteins during the assembly of the head of bacteriophage T4", *Nature*, vol. 227, no. 5259, pages 680–685.

[82] J. G. Lees, A. J. Miles, F. Wien, and B. A. Wallace, 2006. "A reference database for circular dichroism spectroscopy covering fold and secondary structure space", *Bioinformatics*, vol. 22, no. 16, pages 1955–1962.

[83] L. A. Kelley and M. J. E. Sternberg, 2009. "Protein structure prediction on the Web: a case study using the Phyre server", *Nature Protocols*, vol. 4, no. 3, pages 363–371.

[84] W. Kabsch, 1993. "Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants", *Journal of Applied Crystallography*, vol. 26, no. 6, pages 795–800.

[85] S. Bailey, 1994. "The CCP4 suite: programs for protein crystallography", *Acta Crystallographica Section D: Biological Crystallography*, vol. 50, pages 760–763.

[86] R. J. Read, 1986. "Improved Fourier coefficients for maps using phases from partial structures with errors", *Acta Crystallographica Section A: Foundations of Crystallography*, vol. 42, no. 3, pages 140–149.

[87] G. N. Murshudov, A. A. Vagin, A. Lebedev, K. S. Wilson, and E. J. Dodson, 1999. "Efficient anisotropic refinement of macromolecular structures using FFT", *Acta Crystallographica Section D: Biological Crystallography*, vol. 55, no. 1, pages 247–255.

[88] T. A. Jones, J. Y. Zou, S. W. Cowan, and M. Kjeldgaard, 1991. “Improved methods for building protein models in electron density maps and the location of errors in these models”, *Acta Crystallographica Section A: Foundations of Crystallography*, vol. 47, no. 2, pages 110–119.

[89] A. Perrakis, R. Morris, and V. S. Lamzin, 1999. “Automated protein model building combined with iterative structure refinement”, *Nature Structural Biology*, vol. 6, pages 458–463.

[90] J. Painter and E. A. Merritt, 2006. “TLSMD web server for the generation of multi-group TLS models”, *Journal of Applied Crystallography*, vol. 39, no. 1, pages 109–111.

[91] V. A. Marlow and M. Wills, 2009. “Structure and function studies on acetolactate decarboxylase: synthesis of potential substrates and substrate analogues”, MSc Thesis, University of Warwick.

[92] K. B. Sharpless, W. Amberg, Y. L. Bennani, G. A. Crispino, J. Hartung, K. S. Jeong, H. L. Kwong, K. Morikawa, and Z. M. Wang, 1992. “The osmium-catalyzed asymmetric dihydroxylation: a new ligand class and a process improvement”, *The Journal of Organic Chemistry*, vol. 57, no. 10, pages 2768–2771.

[93] M. H. Junttila and O. O. E. Hormi, 2009. “Methanesulfonamide: a cosolvent and a general acid catalyst in Sharpless asymmetric dihydroxylations”, *The Journal of Organic Chemistry*, vol. 74, no. 8, pages 3038–3047.

[94] R. J. Kazlauskas, A. N. E. Weissfloch, A. T. Rappaport, and L. A. Cuccia, 1991. “A rule to predict which enantiomer of a secondary alcohol reacts faster in reactions catalyzed by cholesterol esterase, lipase from *Pseudomonas cepacia*, and lipase from *Candida rugosa*”, *The Journal of Organic Chemistry*, vol. 56, no. 8, pages 2656–2665.

[95] M. Vinogradov, A. Kaplun, M. Vyazmensky, S. Engel, R. Golbik, K. Tittmann, K. Uhlemann, L. Meshalkina, Z. Barak, G. Hübner, and D. M. Chipman, 2005. “Monitoring the acetohydroxy acid synthase reaction and related carboligations by circular dichroism spectroscopy”, *Analytical Biochemistry*, vol. 342, no. 1, pages 126–133.

[96] M. C. Renna, N. Najimudin, L. R. Winik, and S. A. Zahler, 1993. “Regulation of the *Bacillus subtilis alsS, alsD*, and *alsR* genes involved in post-exponential-phase production of acetoin.”, *Journal of Bacteriology*, vol. 175, no. 12, pages 3863–3875.

[97] W. L. Nicholson, 2008. “The *Bacillus subtilis ydjL (bdhA)* gene encodes acetoin reductase/2, 3-butanediol dehydrogenase”, *Applied and Environmental Microbiology*, vol. 74, no. 22, pages 6832–6838.

[98] R. R. de Oliveira and W. L. Nicholson, 2013. “The LysR-type transcriptional regulator (LTTR) AlsR indirectly regulates expression of the *Bacillus subtilis bdhA* gene encoding 2, 3-butanediol dehydrogenase”, *Applied Microbiology and Biotechnology*, pages 1–10.

[99] The UniProt Consortium, 2012. “Reorganizing the protein space at the Universal Protein Resource (UniProt)”, *Nucleic Acids Research*, vol. 40, pages D71–D75.

[100] C. E. Dow, V. A. Marlow, and V. Fülöp, 2011. “Mutagenesis of acetolactate decarboxylase”, MSc Thesis, University of Warwick.

[101] F. W. Studier and B. A. Moffatt, 1986. “Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes.”, *Journal of Molecular Biology*, vol. 189, no. 1, pages 113.

[102] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel, and A. Bairoch, 2005. “Protein identification and analysis tools on the ExPASy server”, *The Proteomics Protocols Handbook*, pages 571–607.

[103] L. A. Compton and W. C. Johnson Jr, 1986. “Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication”, *Analytical Biochemistry*, vol. 155, no. 1, pages 155–167.

[104] P. Manavalan and W. C. Johnson Jr, 1987. “Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra”, *Analytical Biochemistry*, vol. 167, no. 1, pages 76–85.

[105] N. Sreerama and R. W. Woody, 2000. “Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set”, *Analytical Biochemistry*, vol. 287, no. 2, pages 252–260.

[106] A. Lobley, L. Whitmore, and B. A. Wallace, 2002. “DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra”, *Bioinformatics*, vol. 18, no. 1, pages 211–212.

[107] L. Whitmore and B. A. Wallace, 2007. “Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases”, *Biopolymers*, vol. 89, no. 5, pages 392–400.

[108] L. Whitmore and B. A. Wallace, 2004. “DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data”, *Nucleic Acids Research*, vol. 32, no. suppl 2, pages W668–W673.

[109] O. Voges and B. Proskauer, 1898. “Beitrag zur ernährungsphysiologie und zur differentialdiagnose der bakterien der hämorrhagischen septicämie”, *Medical Microbiology and Immunology*, vol. 28, no. 1, pages 20–32.

[110] A. Harden, “On Voges and Proskauer’s reaction for certain bacteria”, *Proceedings of the Royal Society B: Biological Sciences*, vol. 77, 1906, pp. 424–425.

[111] A. Harden and D. Norris, 1911. “The diacetyl reaction for proteins”, *The Journal of Physiology*, vol. 42, no. 4, pages 332–336.

[112] M. M. Barritt, 1936. “The intensification of the Voges-Proskauer reaction by the addition of α-naphthol”, *The Journal of Pathology and Bacteriology*, vol. 42, no. 2, pages 441–454.

[113] T. Nishimura, H. Toku, T. Ueno, S. Shibamoto, and T. Imai, 1972. “Mechanism of Voges-Proskauer reaction”, *Chemistry Letters*, vol. 1, no. 8, pages 649–652.

[114] T. Nishimura, C. Yamazaki, T. Ueno, S. Kitajima, and K. Ishige, 1972. “Studies on the Voges-Proskauer reaction. II. The structure of a pigment from the diacetyl reaction of 1-benzyl-1-methylguanidine”, *Bulletin of the Chemical Society of Japan*, vol. 45, no. 6, pages 1782–1785.

[115] M. I. Dolin and I. C. Gunsalus, 1951. “Pyruvic acid metabolism II.: an acetoin-forming enzyme system in *Streptococcus faecalis*”, *Journal of Bacteriology*, vol. 62, no. 2, pages 199.

[116] E. M. Powers and T. G. Latt, 1977. “Simplified 48-hour IMVic test: an agar plate method.”, *Applied and Environmental Microbiology*, vol. 34, no. 3, pages 274–279.

[117] T. Inoue, 2009. “Diacetyl in Fermented Foods and Beverages”, *Journal of the Institute of Brewing*, vol. 115, no. 2, pages 159–159, ISSN: 2050-0416.

[118] V. A. Marlow and V. Fülöp, 2009. “Structure and function studies on acetolactate decarboxylase: co-crystallisation and kinetic analysis”, MSc Thesis, University of Warwick.

[119] J. Hiratake, 2005. “Enzyme inhibitors as chemical tools to study enzyme catalysis: rational design, synthesis, and applications”, *The Chemical Record*, vol. 5, no. 4, pages 209–228.

[120] N. I. Burzlaff, P. J. Rutledge, I. J. Clifton, C. M. H. Hensgens, M. Pickford, R. M. Adlington, P. L. Roach, and J. E. Baldwin, 1999. “The reaction cycle of isopenicillin N synthase observed by X-ray diffraction”, *Nature*, vol. 401, no. 6754, pages 721–724.

[121] P. L. Roach, I. J. Clifton, C. M. H. Hensgens, N. Shibata, C. J. Schofield, J. Hajdu, and J. E. Baldwin, 1997. “Structure of isopenicillin N synthase complexed with substrate and the mechanism of penicillin formation”, *Nature*, vol. 387, no. 6635, pages 827–829.

[122] T. Bergfors, 2003. “Seeds to crystals”, *Journal of Structural Biology*, vol. 142, no. 1, pages 66–76.

[123] E. A. Stura and I. A. Wilson, 1990. “Analytical and production seeding techniques”, *Methods*, vol. 1, no. 1, pages 38–49.

[124] D. W. J. Cruickshank, 1999. “Remarks about protein structure precision”, *Acta Crystallographica Section D: Biological Crystallography*, vol. 55, no. 3, pages 583–601.

[125] D. H. G. Crout, E. R. Lee, and D. P. J. Pearson, 1990. “A new synthesis of methyl α-acetolactate based on thiazolium chemistry and modelled on the enzymatic synthesis of α-acetolactate catalysed by acetolactate synthase”, *Journal of the Chemical Society, Chemical Communications*, no. 4, pages 331–333.

[126] D. R. Hall, G. A. Leonard, C. D. Reed, C. I. Watt, A. Berry, and W. N. Hunter, 1999. “The crystal structure of *Escherichia coli* class II fructose-1, 6-bisphosphate aldolase in complex with phosphoglycolohydroxamate reveals details of mechanism and specificity”, *Journal of Molecular Biology*, vol. 287, no. 2, pages 383–394.

[127] K. Okrasa, C. Levy, B. Hauer, N. Baudendistel, D. Leys, and J. Micklefield, 2008. “Structure and mechanism of an unusual malonate decarboxylase and related racemases”, *Chemistry-A European Journal*, vol. 14, no. 22, pages 6609–6613.

[128] K. Okrasa, C. Levy, M. Wilding, M. Goodall, N. Baudendistel, B. Hauer, D. Leys, and J. Micklefield, 2009. “Structure-guided directed evolution of alkenyl and arylmalonate decarboxylases”, *Angewandte Chemie*, vol. 121, no. 41, pages 7827–7830.

[129] J. H. Hurley, A. M. Dean, D. E. Koshland Jr, and R. M. Stroud, 1991. “Catalytic mechanism of $NADP^+$-dependent isocitrate dehydrogenase: implications from the structures of magnesium-isocitrate and $NADP^+$ complexes”, *Biochemistry*, vol. 30, no. 35, pages 8671–8678.

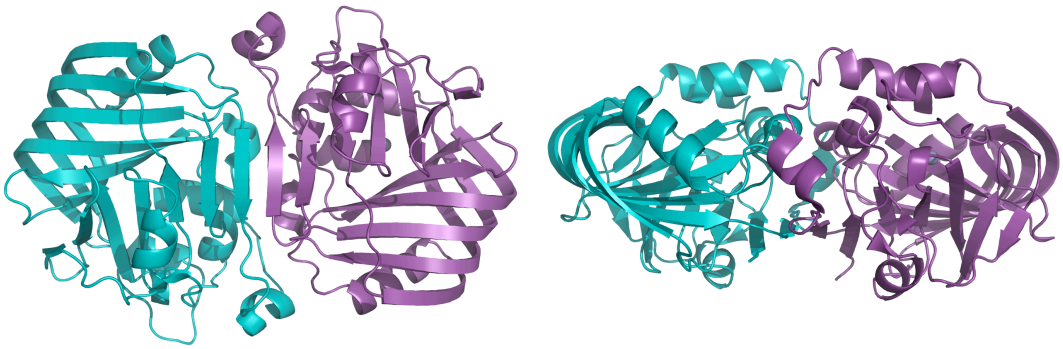
[130] S. S. Pang, R. G. Duggleby, R. L. Schowen, and L. W. Guddat, 2004. “The crystal structures of *Klebsiella pneumoniae* acetolactate synthase with enzyme-bound cofactor and with an unusual intermediate”, *Journal of Biological Chemistry*, vol. 279, no. 3, pages 2242–2253.

[131] R. Wohlgemuth, 2010. “Asymmetric biocatalysis with microbial enzymes and cells”, *Current Opinion in Microbiology*, vol. 13, no. 3, pages 283–292.

Appendix A

S. aureus ALDC structure

Figure A.1: Structure of *S. aureus* ALDC



(a) The structure of S. aureus ALDC at 2 Å resolution. Monomers are shown in cyan and purple, the two N-terminal domains form a β sheet at the dimer interface. (b) The dimer viewed from a 90° rotation shows the β barrel from the C-terminal domain. PDB accession code 1XV2

Appendix B

NMR spectra

This appendix contains the NMR spectra obtained for all compounds synthesised in this project as described in Chapter 2 Methods and experimental.

Figure B.1: Proton NMR of (*threo*)-methyl-2,3-dihydroxy-2-methylbutanoate (*threo*)-10

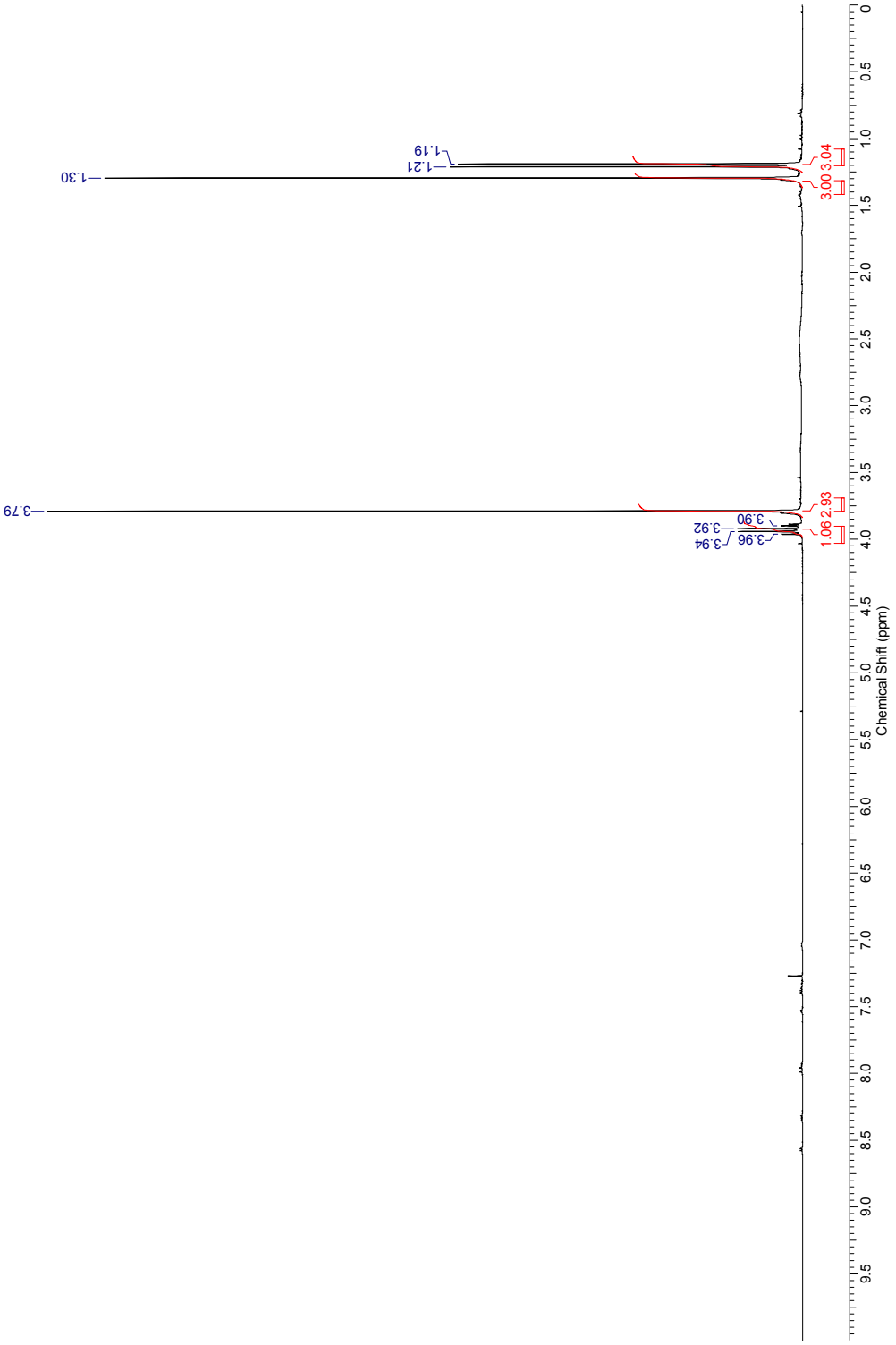


Figure B.2: Carbon NMR of (*threo*)-methyl-2,3-dihydroxy-2-methylbutanoate (*threo*)-10

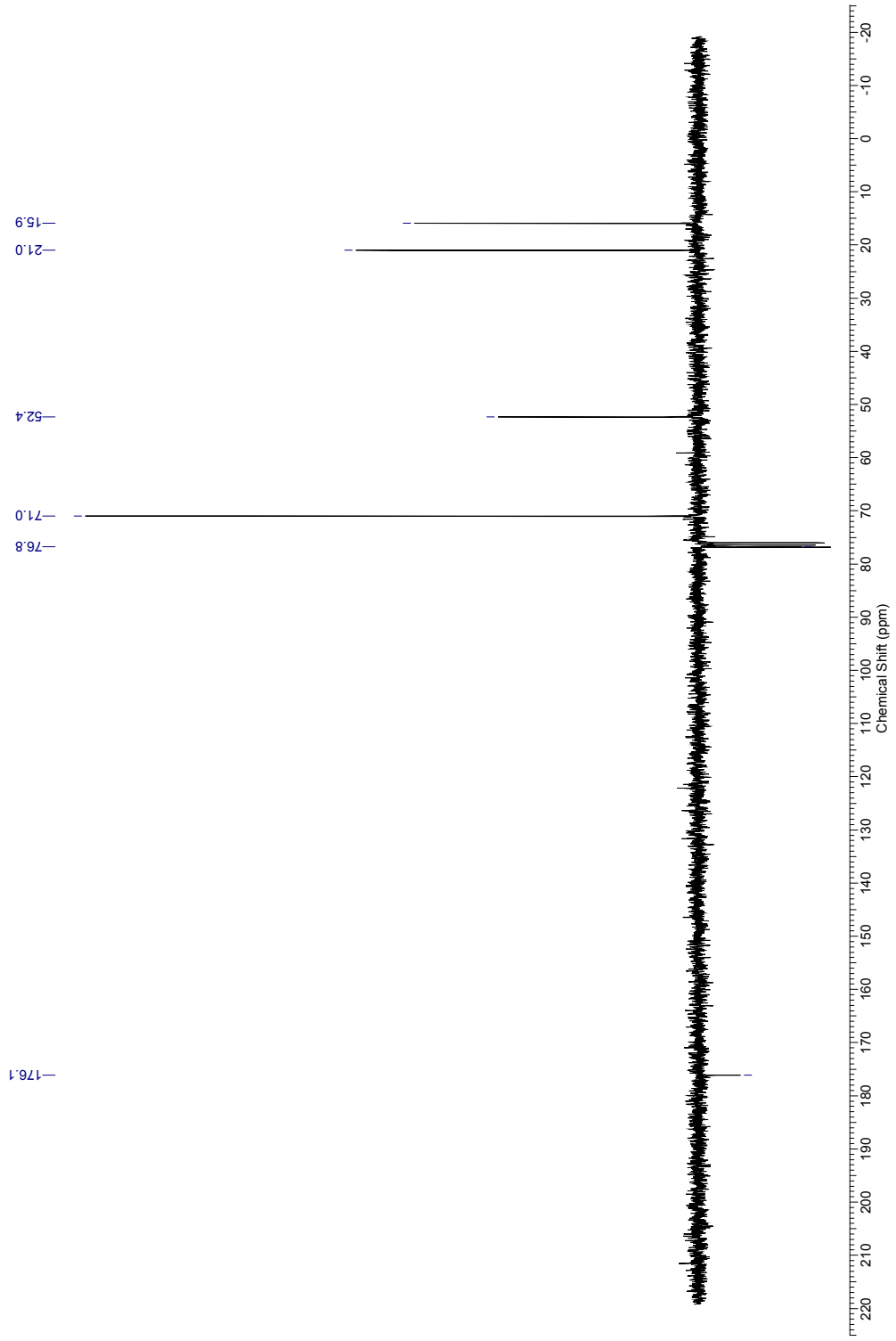


Figure B.3: Proton NMR of (*erythro*)-methyl-2,3-dihydroxy-2-methylbutanoate (*erythro*)-10

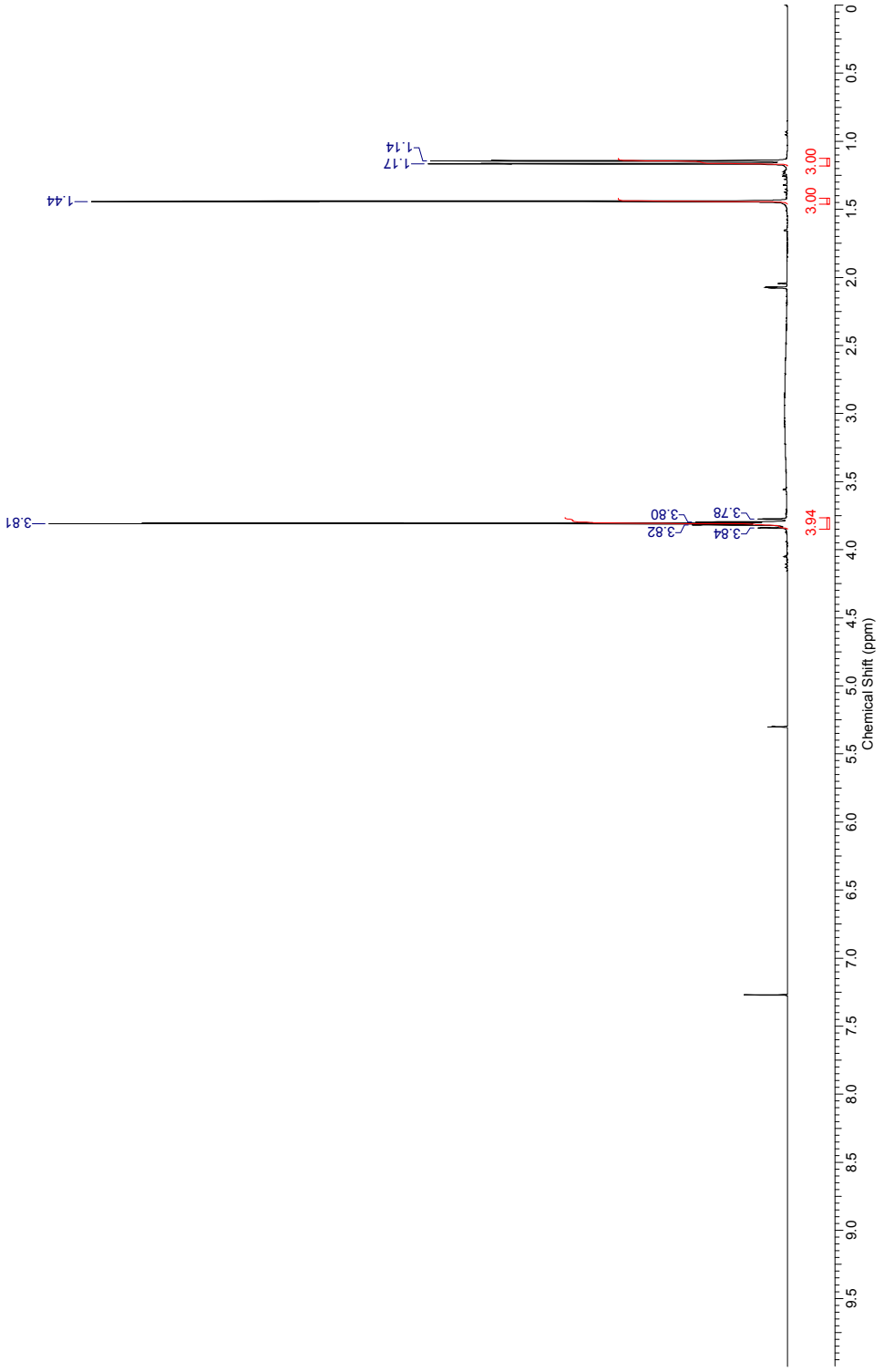


Figure B.4: Carbon NMR of (*erythro*)-methyl-2,3-dihydroxy-2-methylbutanoate (*erythro*)-10

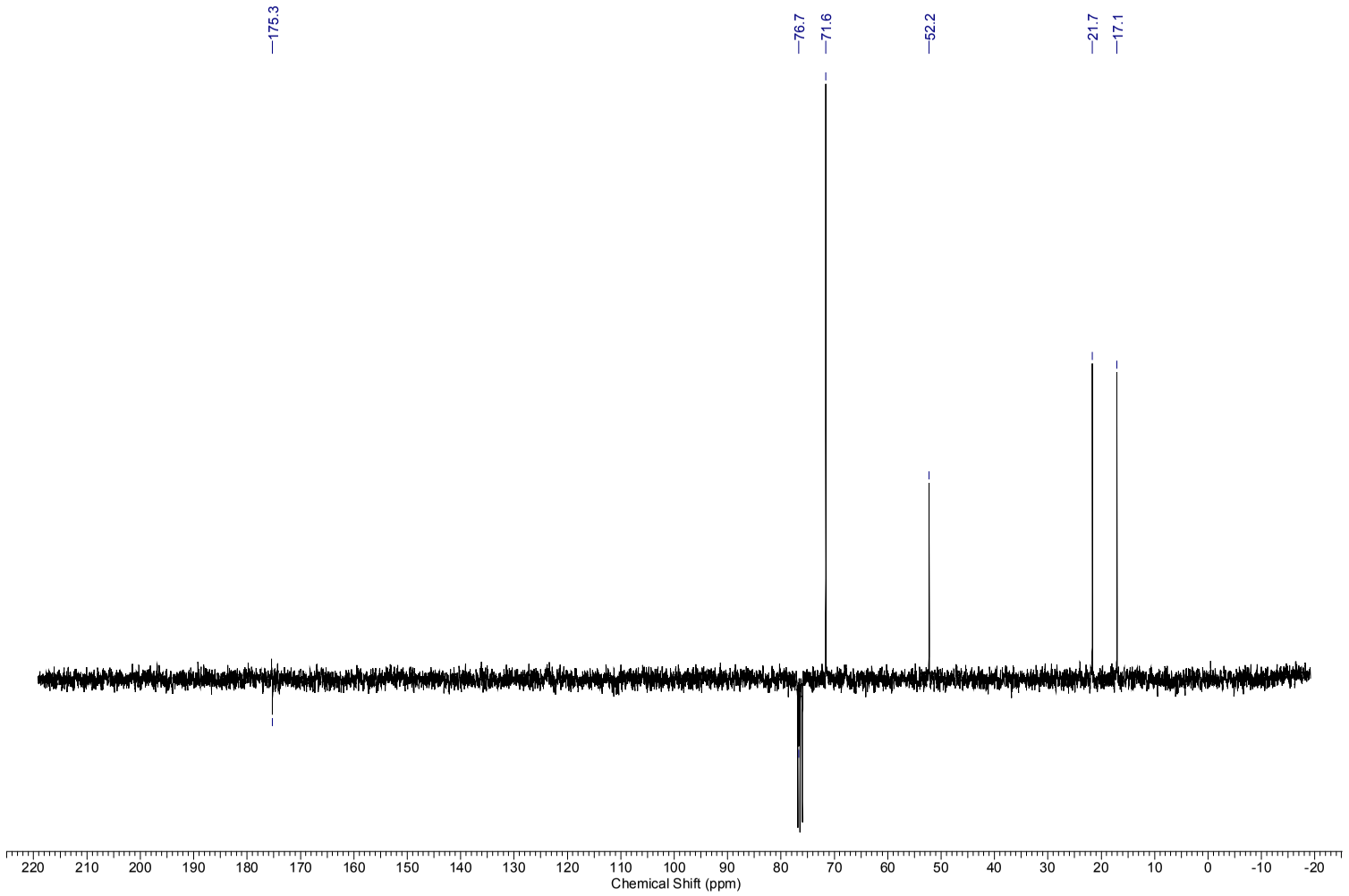


Figure B.5: Proton NMR of (*threo*)-2,3-dihydroxy-2-methylbutanoic acid (*threo*)-11

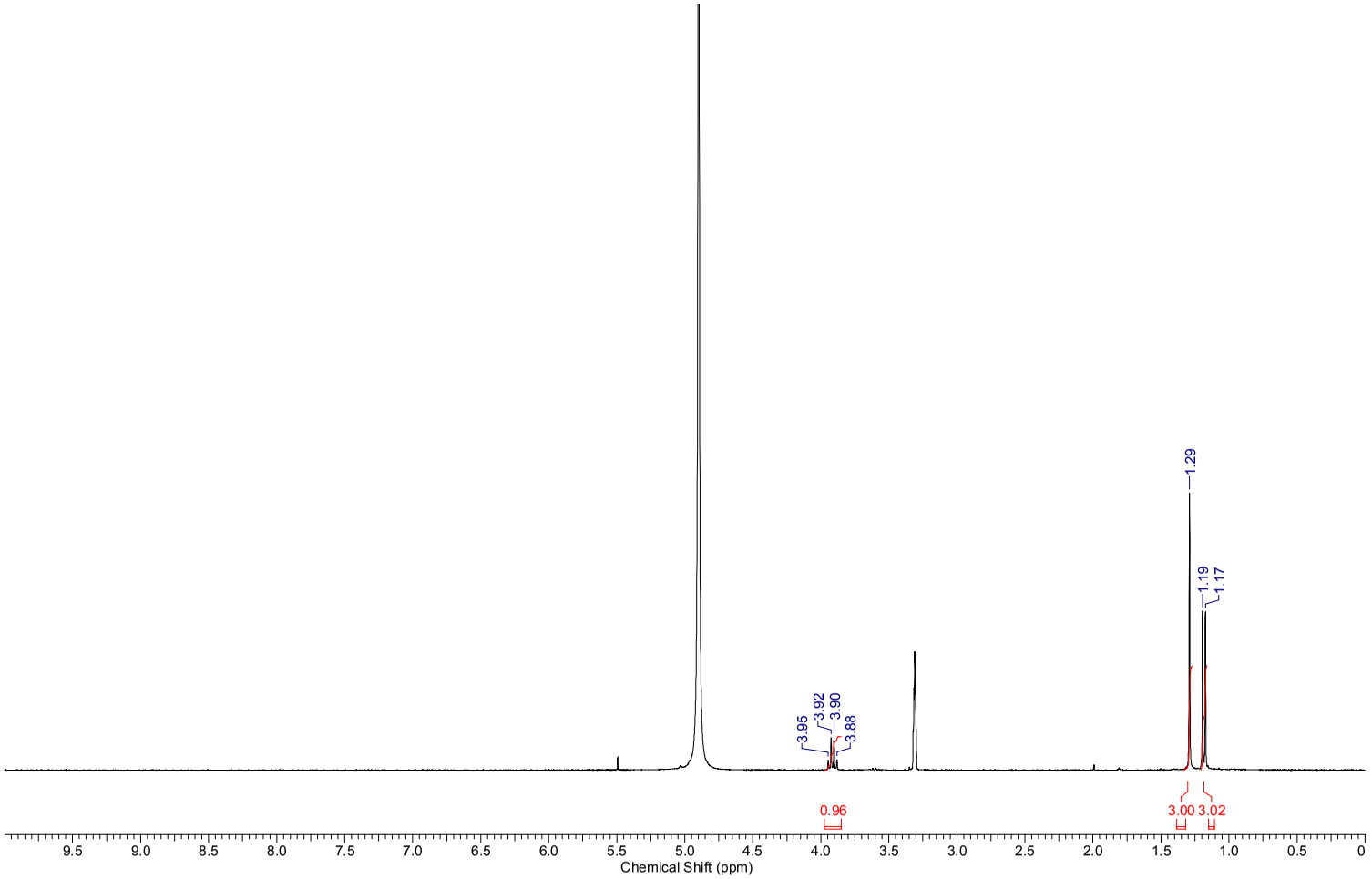


Figure B.6: Carbon NMR of (*threo*)-2,3-dihydroxy-2-methylbutanoic acid (*threo*)-11

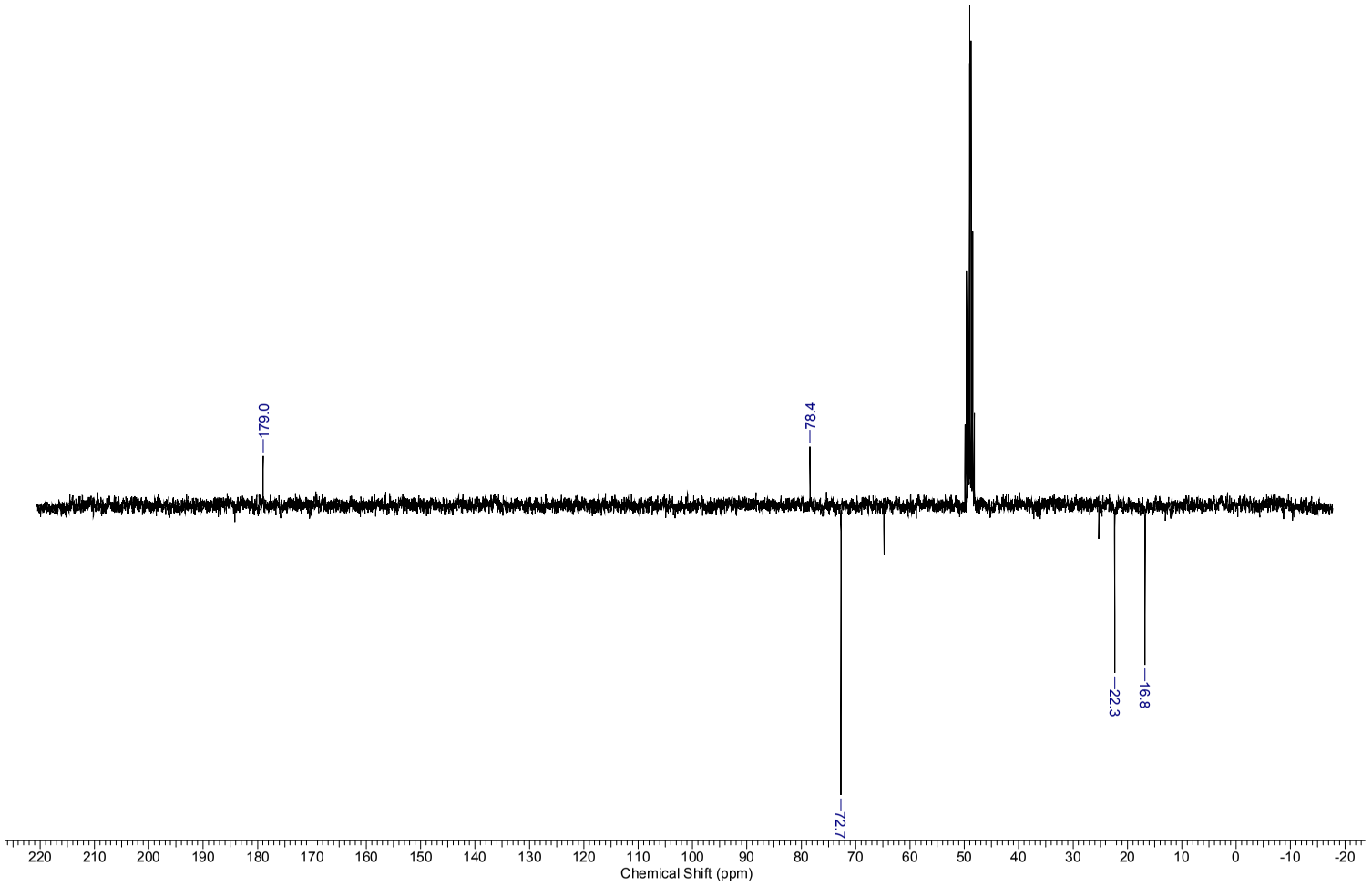


Figure B.7: Proton NMR of (*erythro*)-2,3-dihydroxy-2-methylbutanoic acid (*erythro*)-11

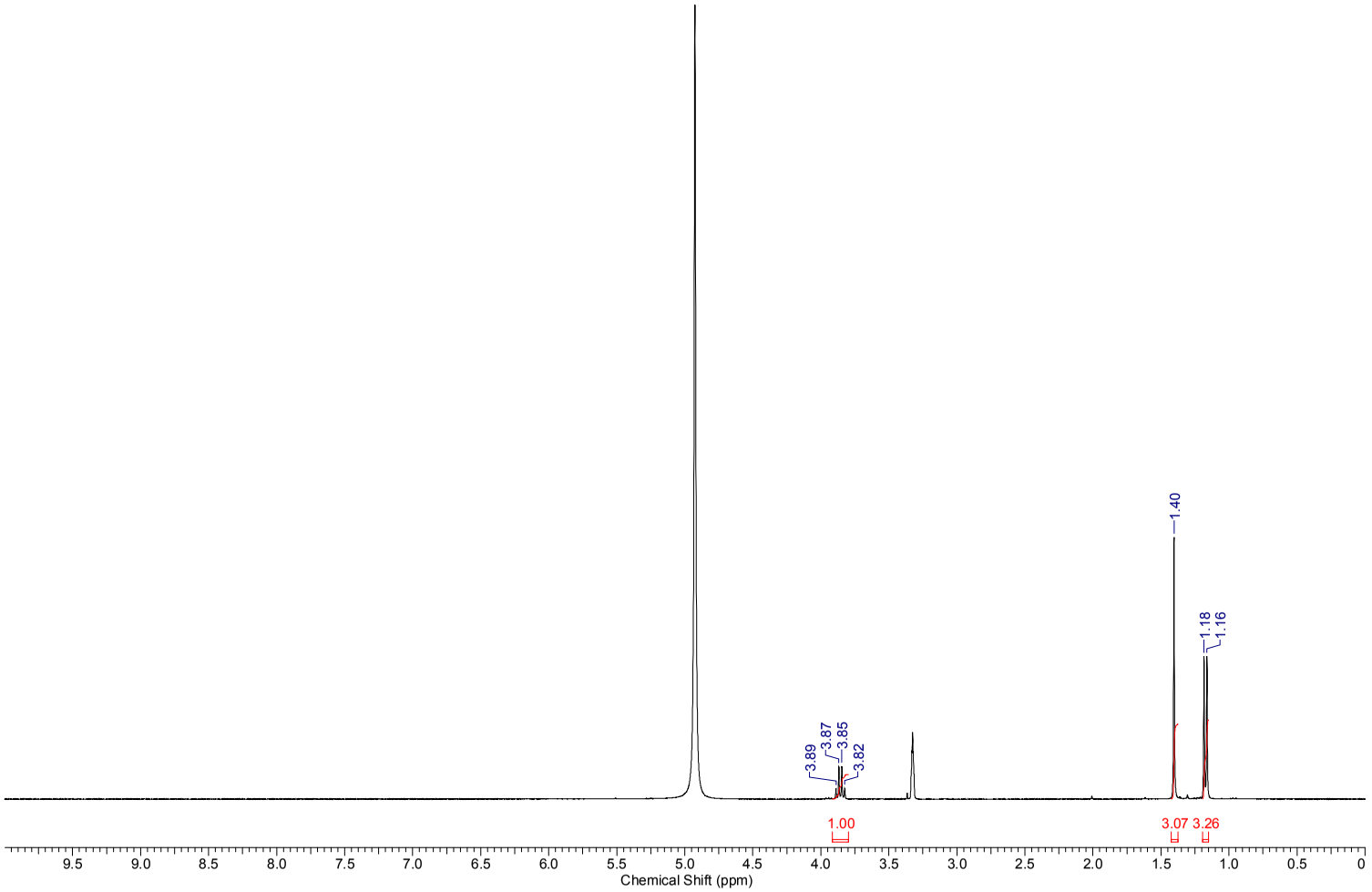
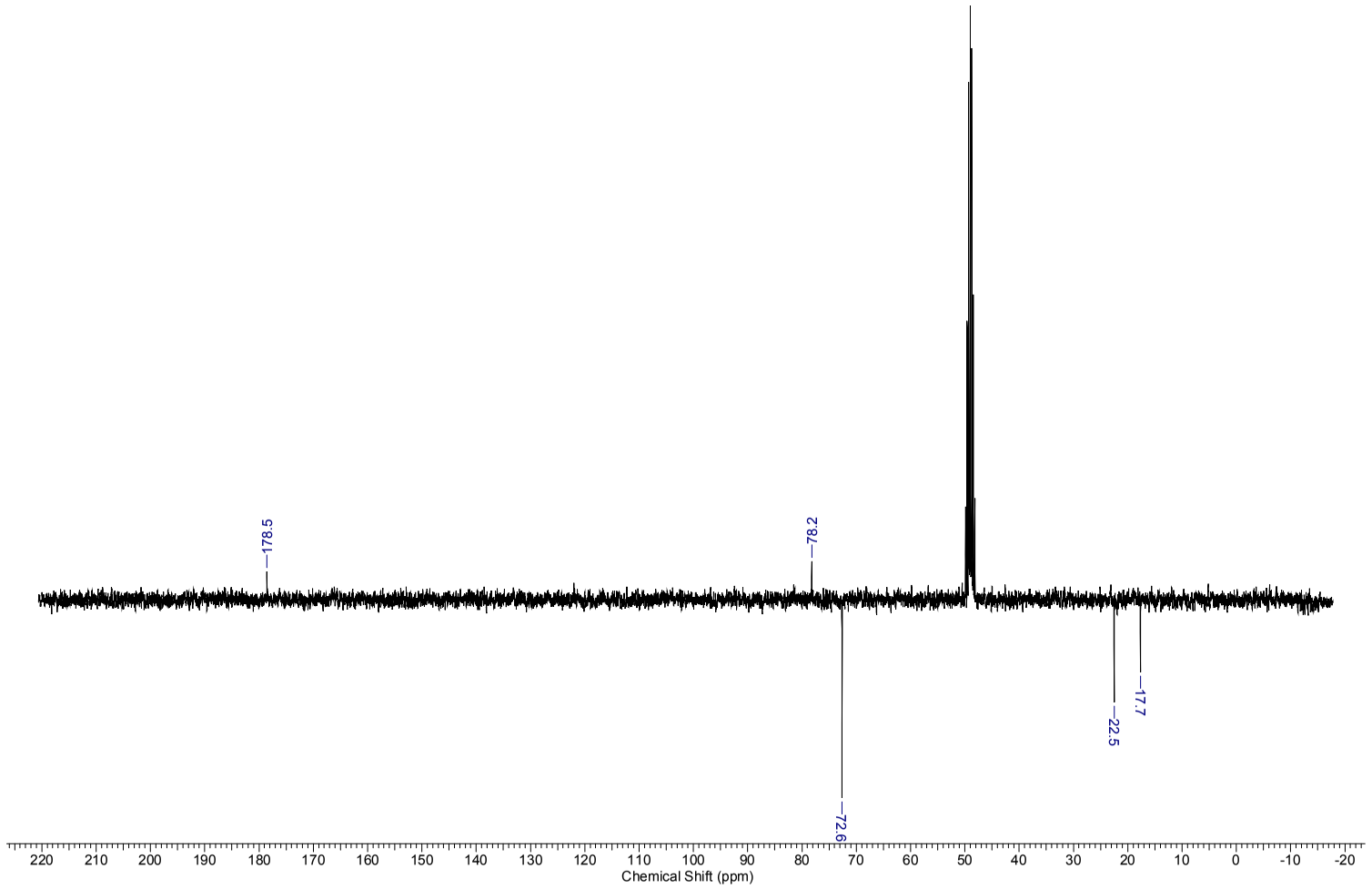


Figure B.8: Carbon NMR of (*erythro*)-2,3-dihydroxy-2-methylbutanoic acid (*erythro*)-11



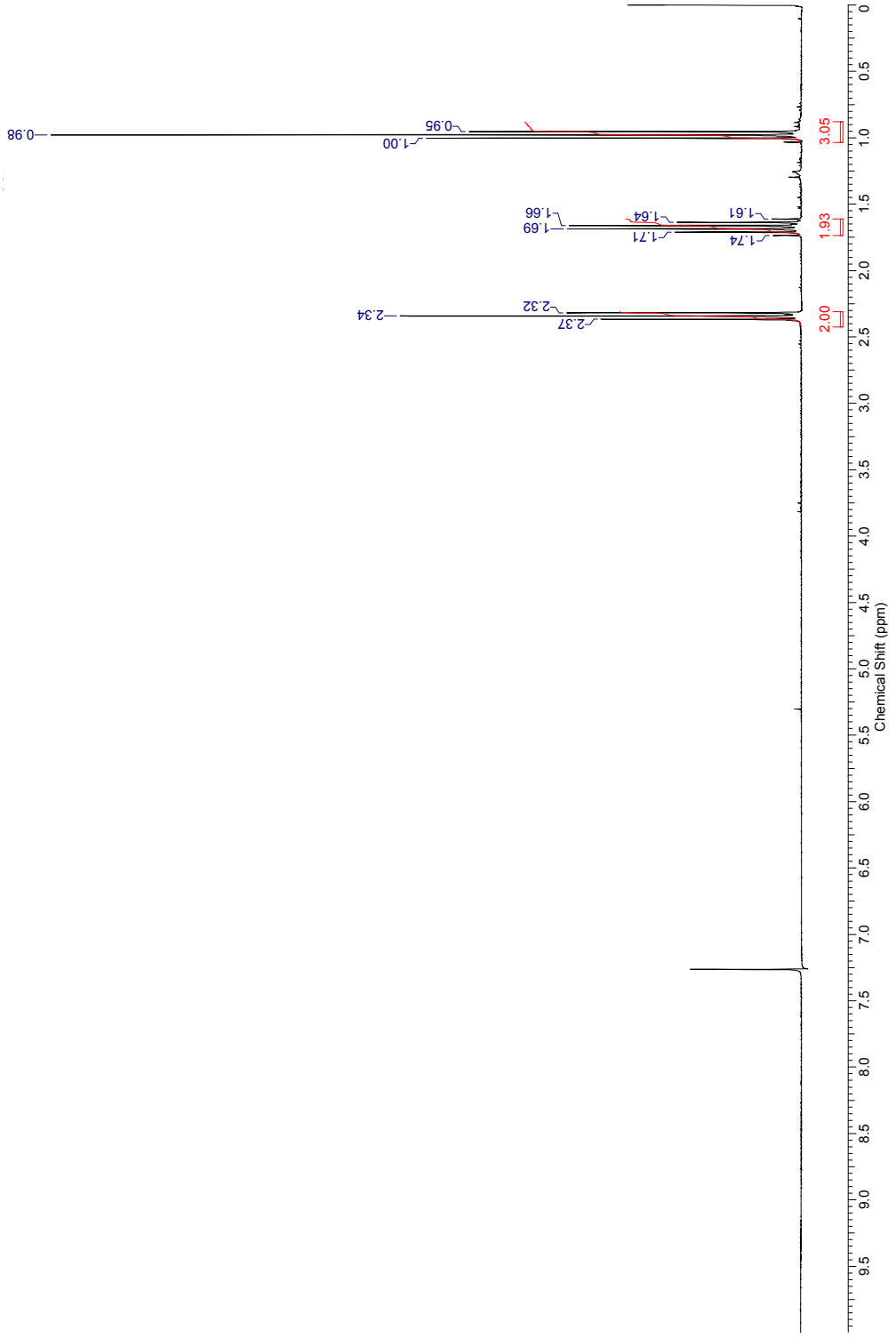


Figure B.9: Proton NMR of butyric acid 14

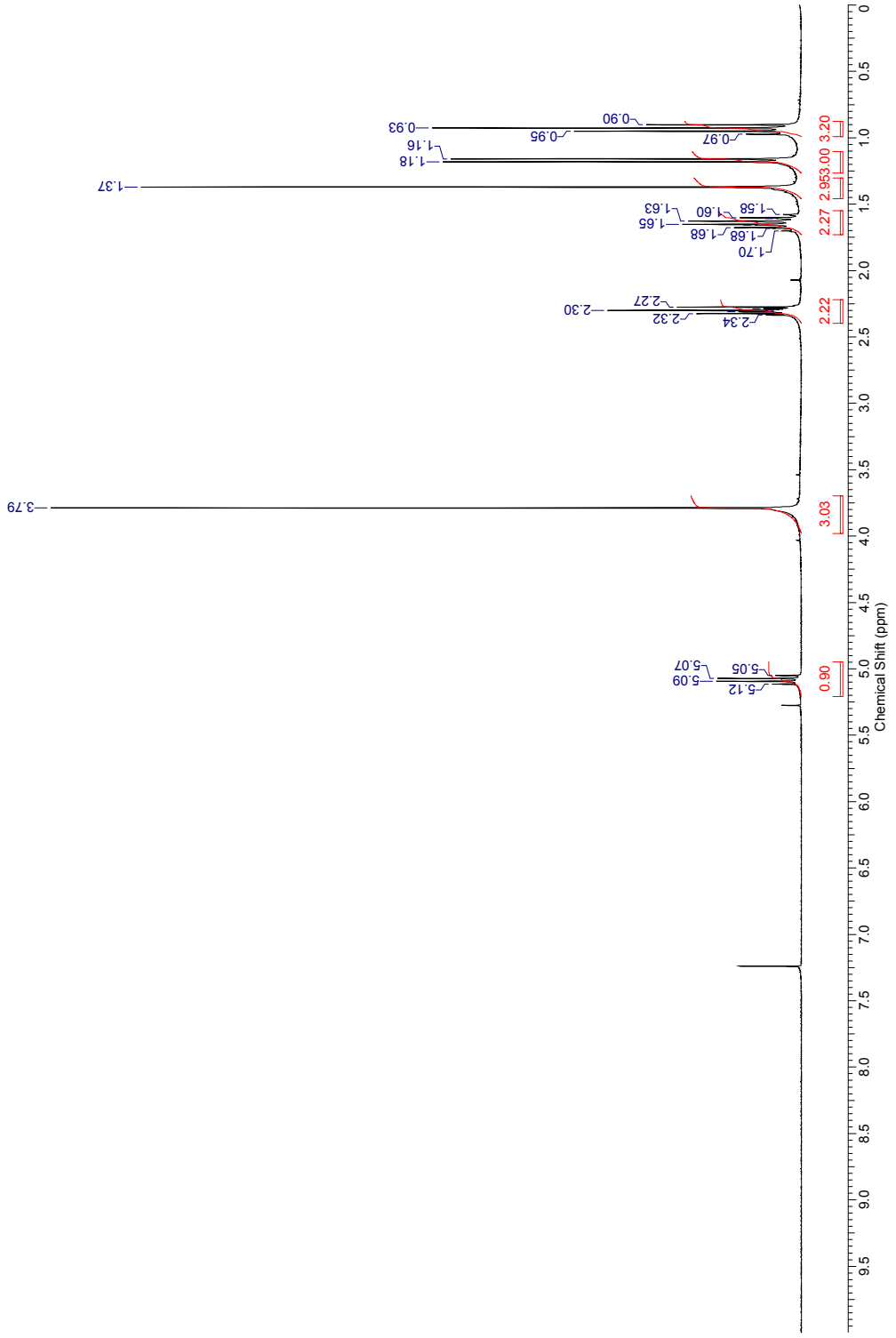
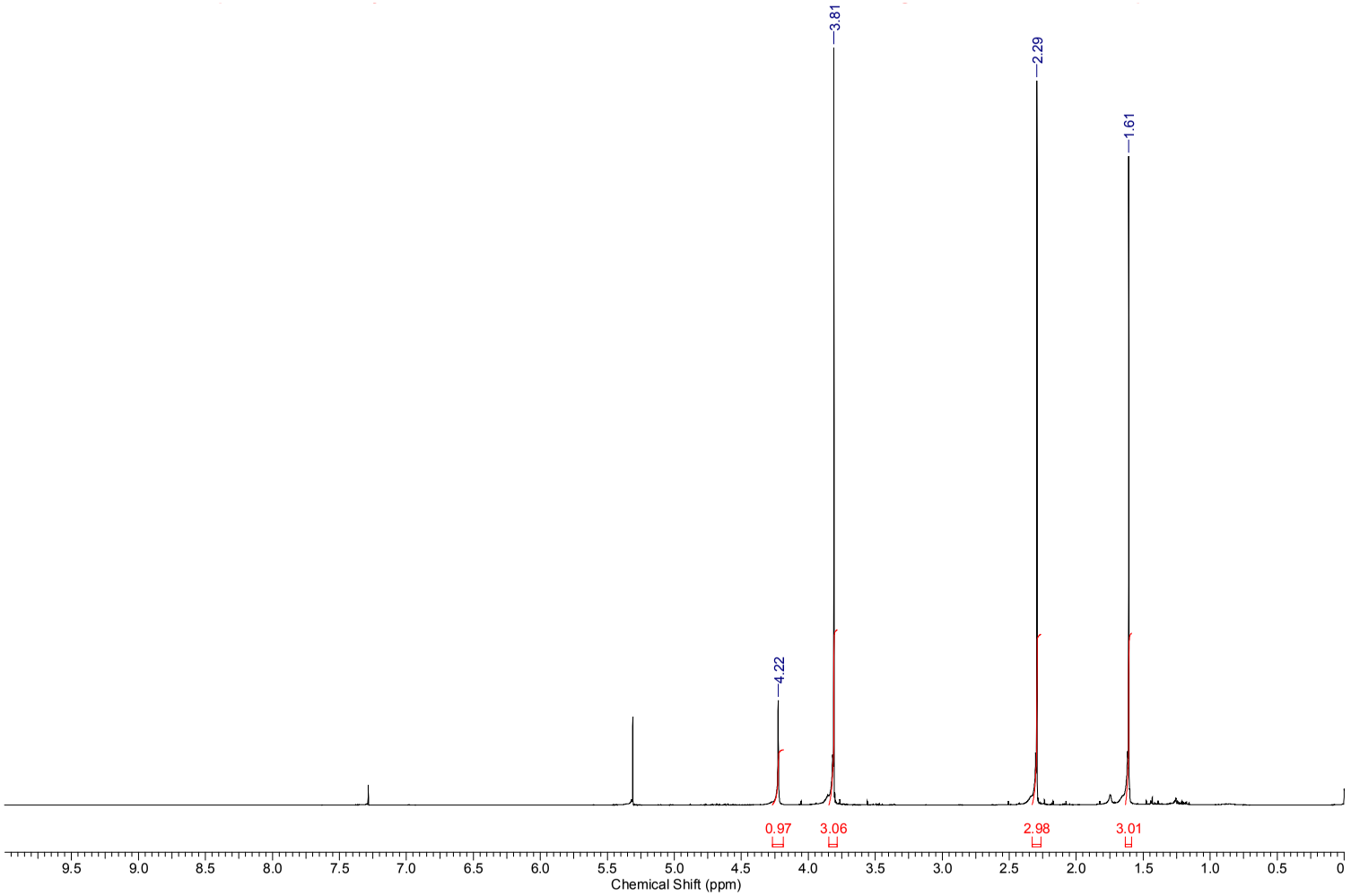


Figure B.10: Proton NMR of methyl-2-hydroxy-2-methyl-3-(butanoyloxy)butanoate 13

Figure B.11: Proton NMR of methyl acetolactate 12



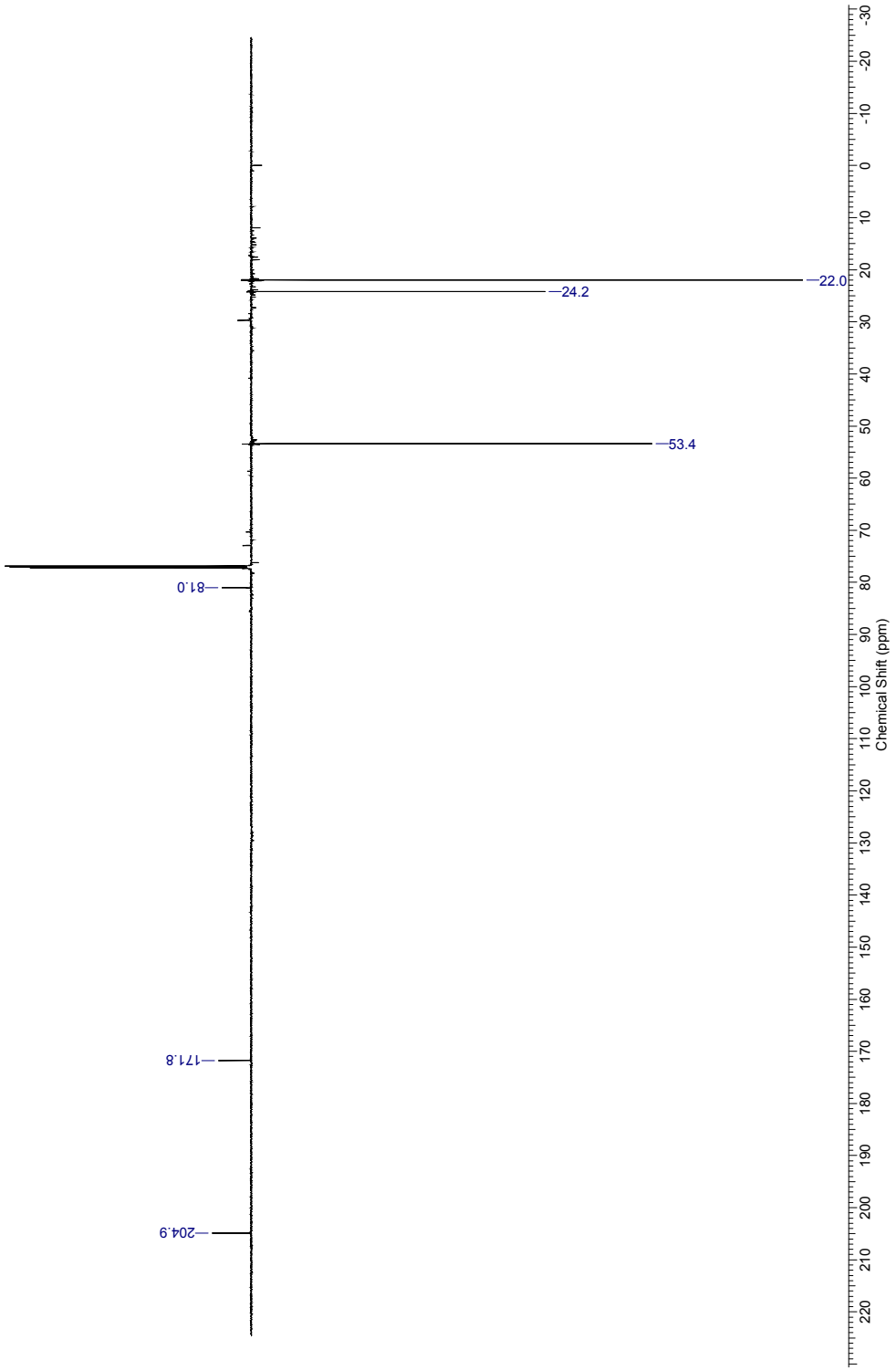


Figure B.12: Carbon NMR of methyl acetolactate 12

Figure B.13: Proton NMR of methyl-2-methyl-2-(methylsulphanylmethoxy)-3-oxobutanoate 15

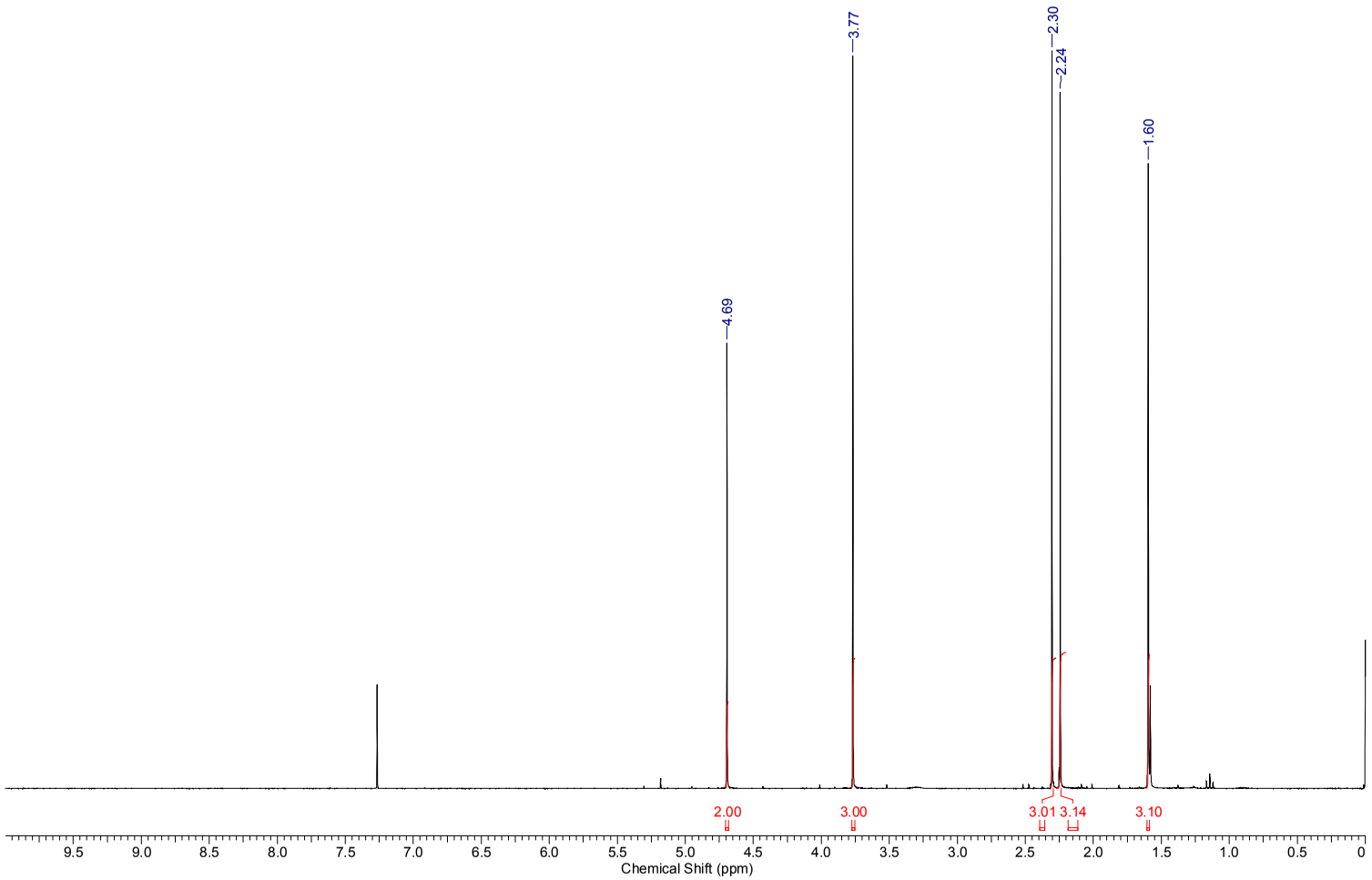
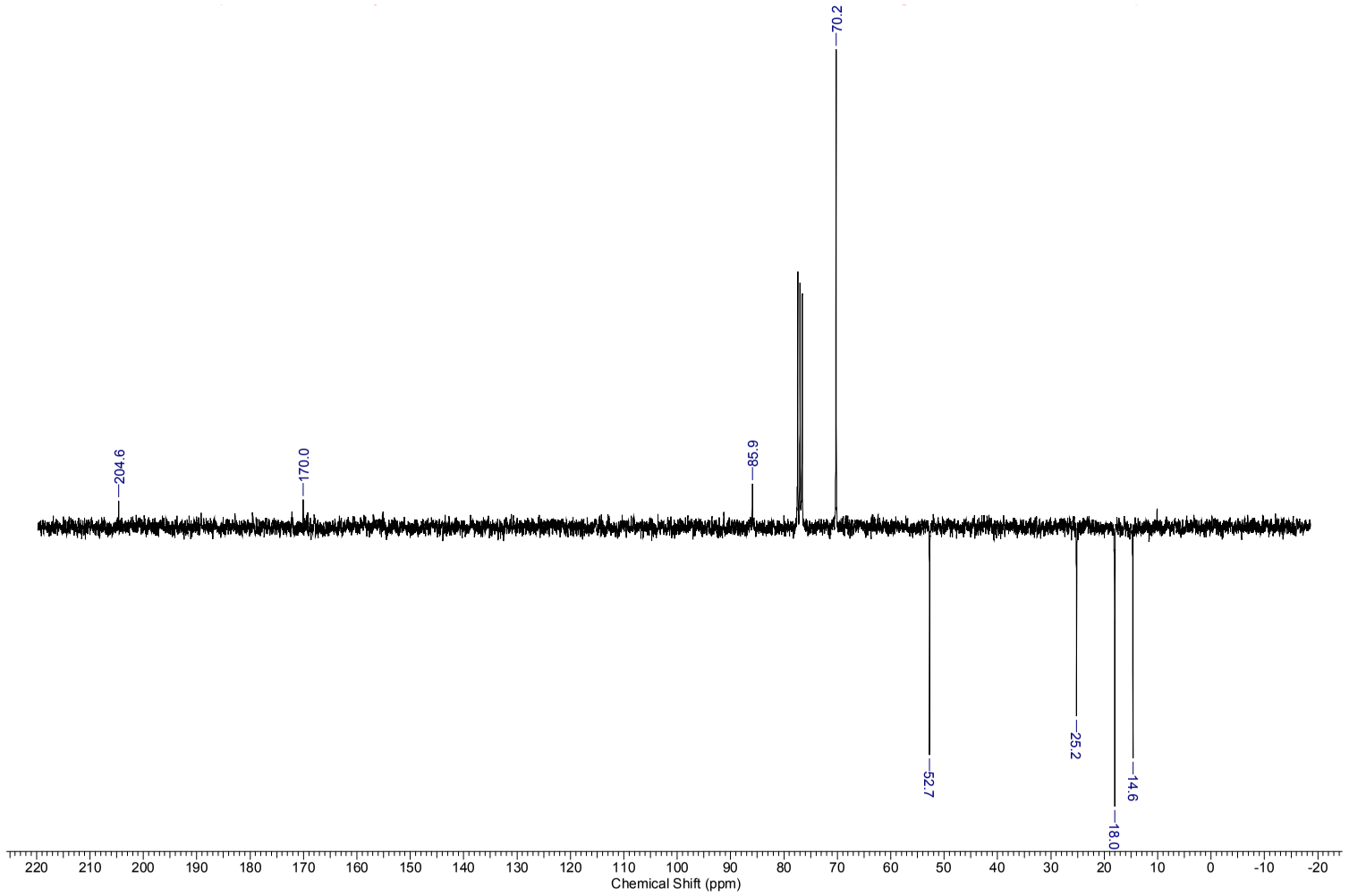


Figure B.14: Carbon NMR of methyl-2-methyl-2-(methylsulphanylmethoxy)-3-oxobutanoate 15



Appendix C

Sigma Aldrich ALDC Protocol

The following protocol can be found online at http://www.sigmaaldrich.com/content/dam/sigma-aldrich/docs/Sigma/General_Information/a-acetolactate_decarboxylase.pdf

SIGMA QUALITY CONTROL TEST PROCEDURE

ProductInformation

Enzymatic Assay of α-ACETOLACTATE DECARBOXYLASE
(EC 4.1.1.5)

PRINCIPLE:

2-Acetolactate $\xrightarrow{\text{α-Acetolactate Decarboxylase}}$ Acetoin + CO_2

CONDITIONS: T = 30°C, pH = 6.0, A_{522nm}, Light path = 1 cm

METHOD: Stopped Spectrophotometric Rate Determination

REAGENTS:

A. 50 mM MES Buffer, pH 6.0 at 30°C
(Prepare 100 ml in deionized water using MES Free Acid, Monohydrate, Sigma Stock No. M-5287. Adjust to pH 6.0 at 30°C with 1 M NaOH.)

B. 15% (w/v) Brij 35 Solution (Brij 35)
(Prepare 2 ml in deionized water using Brij 35 Solution, 30% w/v solution, Sigma Stock No. 430AG-6. Heat to no more than 60°C to dissolve.)

C. 50 mM MES Buffer, pH 6.0 at 30°C with 0.005% (w/v) Brij 35 and 600 mM Sodium Chloride (Buffer C)
(Prepare 500 ml by dissolving 4.9 g of MES, Free Acid, Monohydrate, Sigma Prod. No. M-5287 and 17.5 g of Sodium Chloride, Sigma Prod. No. S-9625 in approximately 450 ml of deionized water. Add 1.7 ml of Reagent B and adjust to pH 6.0 at 30°C with 1 M NaOH.)

D. 10 mM α-Acetolactate Substrate Solution (α-AL)
(Prepare by diluting 0.05 ml of Ethyl 2-acetoxy-2-methylacetoacetate, Aldrich Stock No. 22,039-6 with 1.5 ml of deionized water and 1.5 ml of 1 N NaOH. Stir 20 minutes, Q.S. to 20 ml with Reagent A. Adjust to pH 6.0 with 0.5 M HCl and then Q.S. to 25 ml with Reagent A.

E. Color Reagent Solution (Color Rgt)
(Prepare 250 ml by dissolving 2.5 g of α-Naphthol, Sigma Prod. No. N-1000 and 0.25 g of Creatine, Hydrate, Sigma Prod. No. C-3630 in 1 N NaOH (use a minimum amount). Q.S. to 250 ml with 1 N NaOH.)

Enzymatic Assay of α-ACETOLACTATE DECARBOXYLASE
(EC 4.1.1.5)

REAGENTS: (continued)

F. 2.27 mM Acetoin Stock Solution
(Prepare 100 ml in deionized water using 3-Hydroxy-2-Butanone, Aldrich Stock No. A1,795-1.)

H. α-Acetolactate Decarboxylase Enzyme Solution
(Immediately before use, prepare a solution containing approximately 0.03 unit/ml of α-Acetolactate Decarboxylase in Reagent C.)

PROCEDURE:

Part I

Pipette (in milliliters) the following reagents into suitable containers:

	Test	Enzyme Blank
Reagent D (α-AL)	0.20	0.20
Reagent H (Enz Sol)	0.20	------
Reagent C (Buffer C)	------	------

Incubate for 20 minutes at 30°C in a water bath. Then add:

Reagent E (Color Rgt)	4.50	4.50
Reagent H (Enz Sol)	------	0.20

Incubate for 40 minutes at room temperature. Transfer the solutions to suitable cuvettes and record the A_{522nm} for the Tests and Blanks using a suitable spectrophotometer.

Part II Standard Curve

Pipette (in milliliters) the following reagents into suitable 10 ml plastic tubes:

	Std 1	Std 2	Std 3	Std 4	Std 5	Std Blank
Acetoin	0.020	0.040	0.080	0.120	0.160	------
Milli-Q	0.380	0.360	0.320	0.280	0.240	0.400
Color Solution	4.50	4.50	4.50	4.50	4.50	4.50

Color develops in 40 minutes at room temperature. Transfer the solutions to suitable cuvettes and record the A_{522nm} for the Standards and Standard Blank using a suitable spectrophotometer.

Enzymatic Assay of α-ACETOLACTATE DECARBOXYLASE
(EC 4.1.1.5)

CALCULATIONS:

Standard Curve:

ΔA_{522nm} Standard = A_{522nm} Standard - A_{522nm} Standard Blank

Prepare a standard curve by plotting the ΔA_{522nm} of the Standard versus µmoles of acetoin.

Sample Determination:

ΔA_{522nm} Test = A_{522nm} Test - A_{522nm} Test Blank

Determine the micromoles of acetoin produced using the standard curve.

$$\text{Units/ml enzyme} = \frac{(\mu\text{moles of acetoin produced})(4.9)(df)}{(20)(0.2)}$$

4.9 = Total volume (in milliliters) of assay
df = Dilution factor
20 = Time (in minutes) of assay
0.2 = Volume (in milliliters) of enzyme used

$$\text{Units/mg solid} = \frac{\text{units/ml enzyme}}{\text{mg solid/ml enzyme}}$$

$$\text{Units/mg protein} = \frac{\text{units/ml enzyme}}{\text{mg protein/ml enzyme}}$$

UNIT DEFINITION:

One unit will form one µmole of acetoin from acetolactate per minute at pH 6.0 at 30°C.

FINAL ASSAY CONCENTRATION:

In a 0.40 ml reaction mix, the final concentrations are 5 mM DL-acetolactate, 48 mM MES (approximately) and 0.006 unit α-acetolactate decarboxylase.

REFERENCE:

Stormer, F.C. (1975) *Methods in Enzymology,* XLI, 526-529

Enzymatic Assay of α-ACETOLACTATE DECARBOXYLASE
(EC 4.1.1.5)

NOTES:

1. This assay is based on the cited reference.

2. Where Sigma Product or Stock numbers are specified, equivalent reagents may be substituted.

Appendix D

Mass spectra

This appendix contains the mass spectra obtained for the recombinant proteins that were produced as described in Chapter 2 Methods and experimental.

Figure D.1: Mass spectrum of recombinant wild-type AlsD

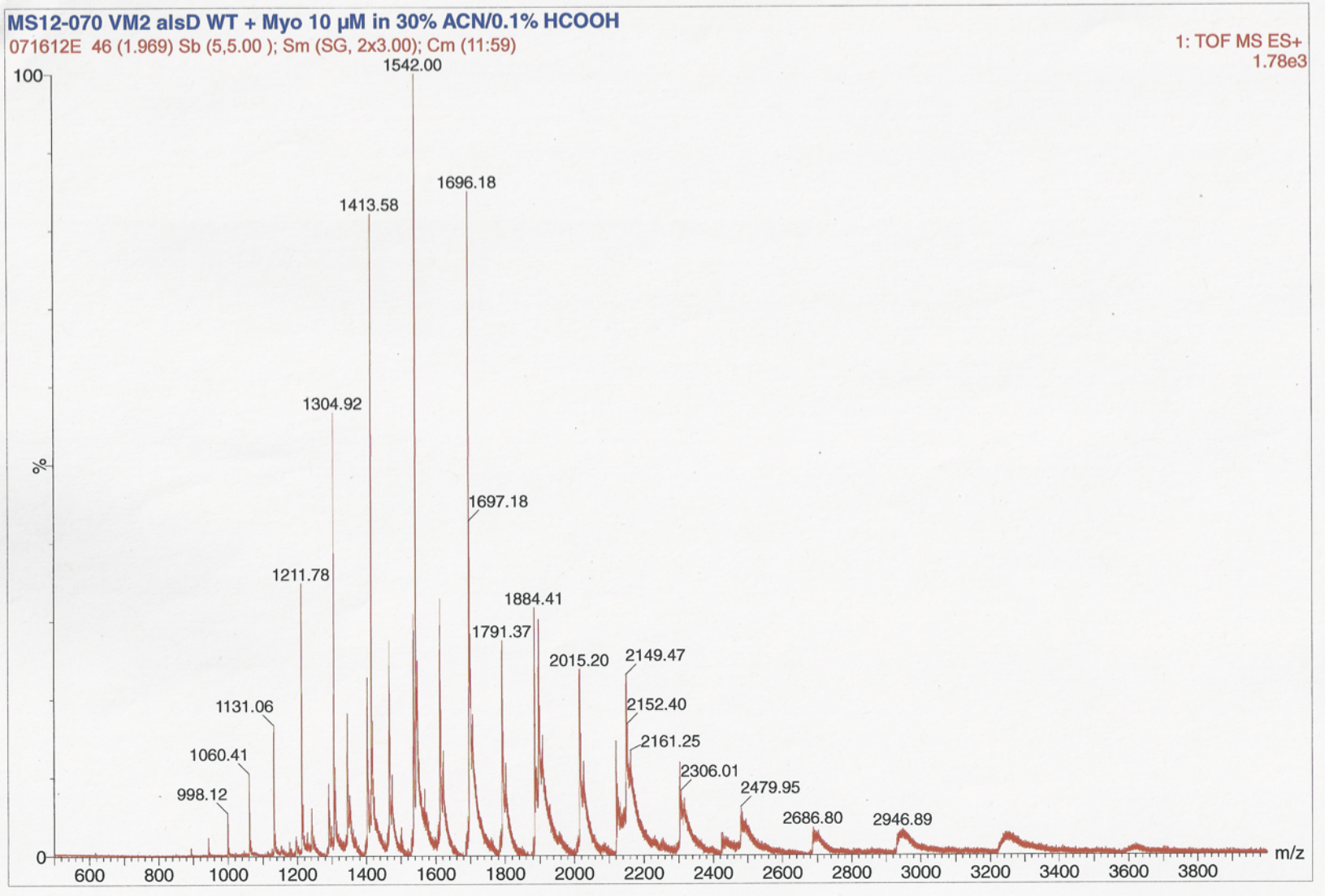


Figure D.2: Deconvoluted mass spectrum of recombinant wild-type AlsD

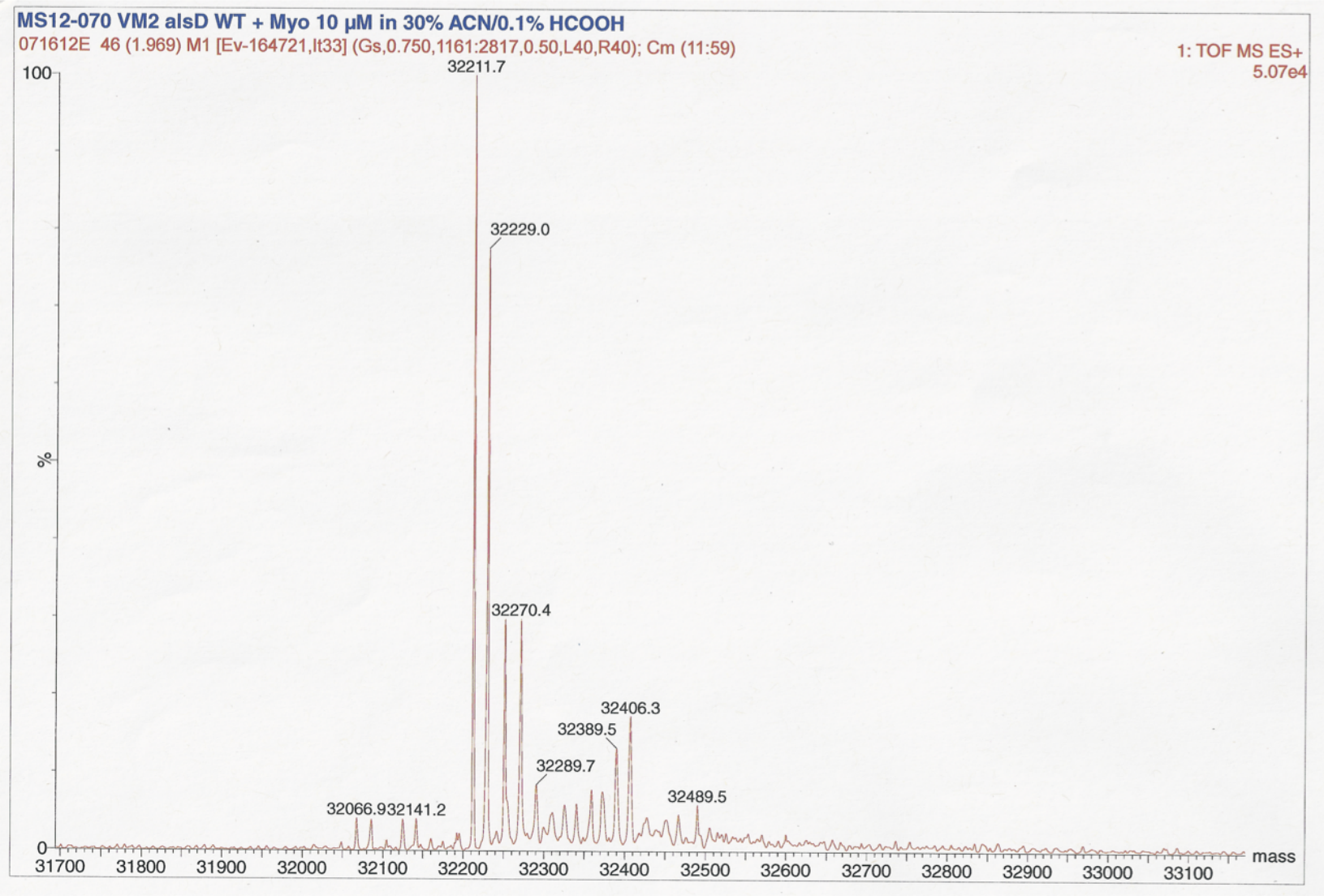


Figure D.3: Mass spectrum of recombinant E62Q AlsD

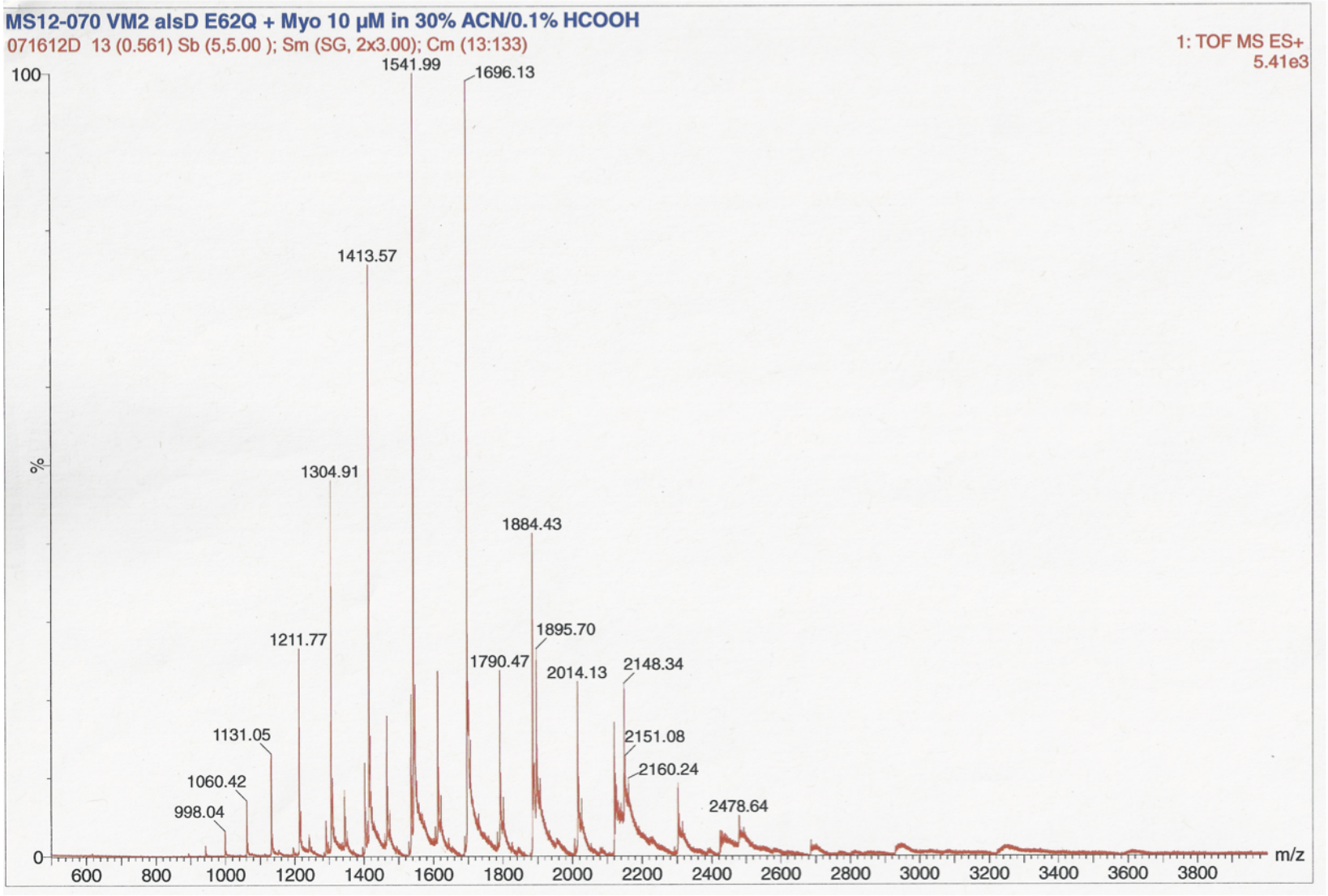


Figure D.4: Deconvoluted mass spectrum of recombinant E62Q AlsD

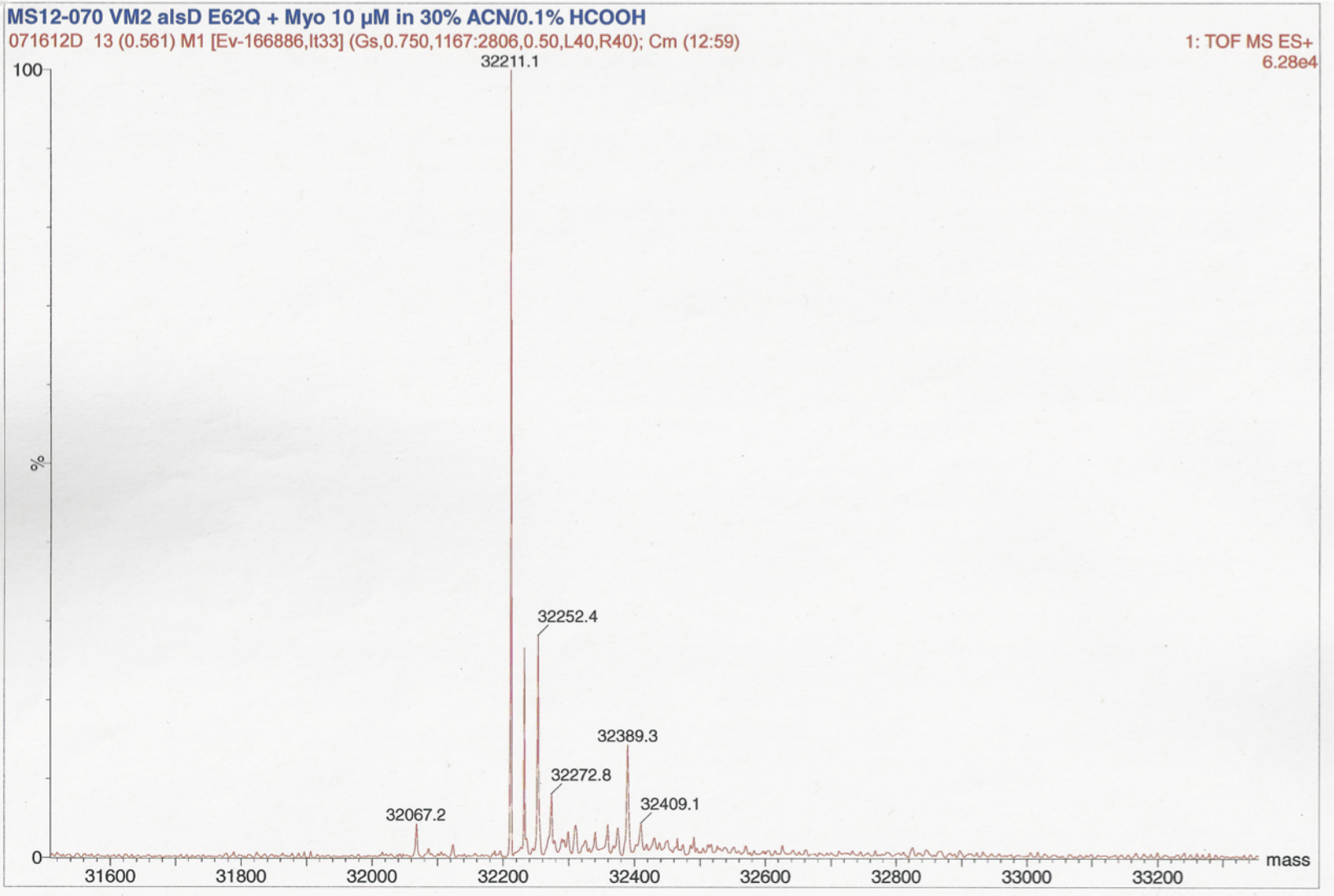


Figure D.5: Mass spectrum of recombinant E62A AlsD

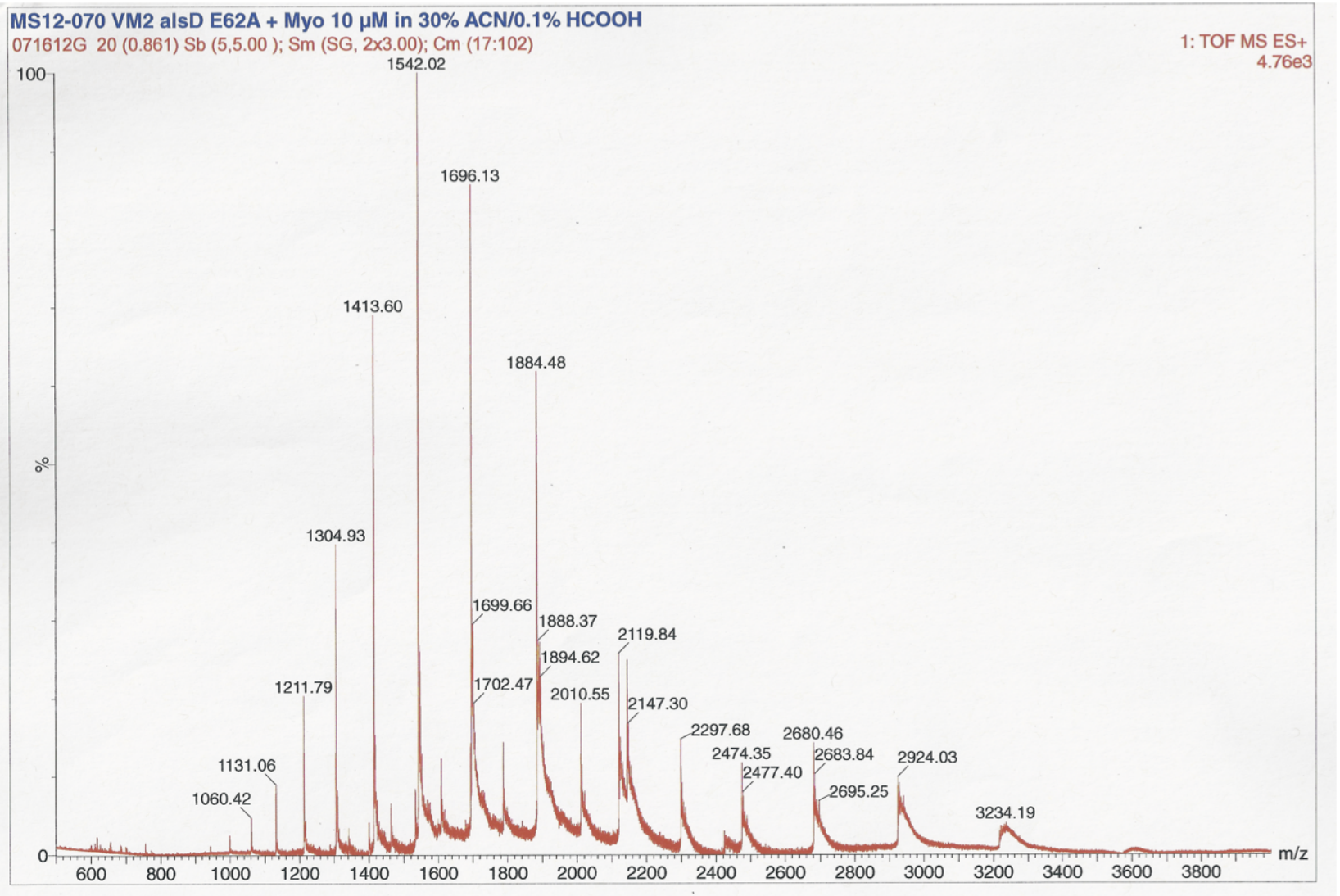


Figure D.6: Deconvoluted mass spectrum of recombinant E62A AlsD

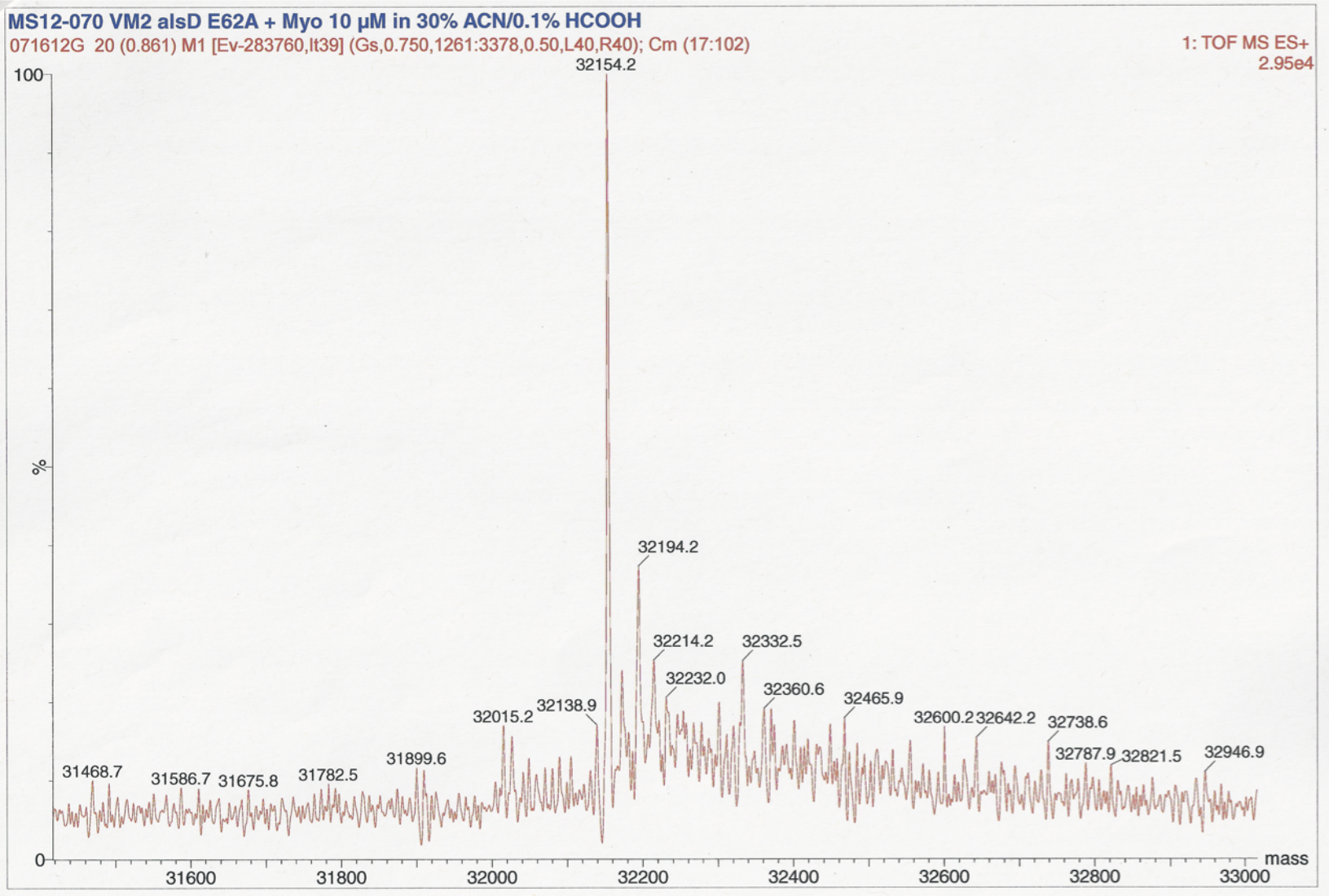


Figure D.7: Mass spectrum of recombinant R142K AlsD

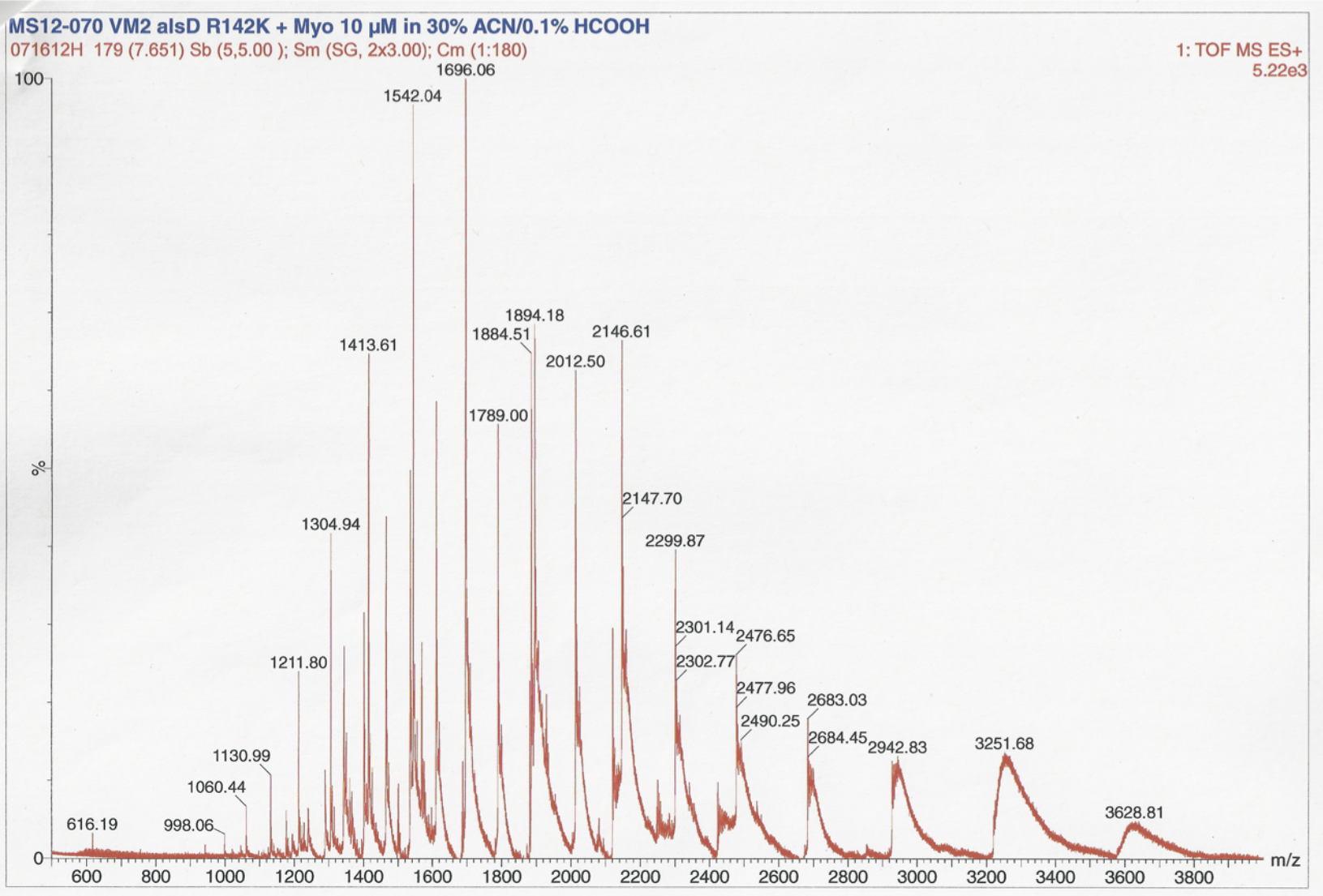


Figure D.8: Deconvoluted mass spectrum of recombinant R142K AlsD

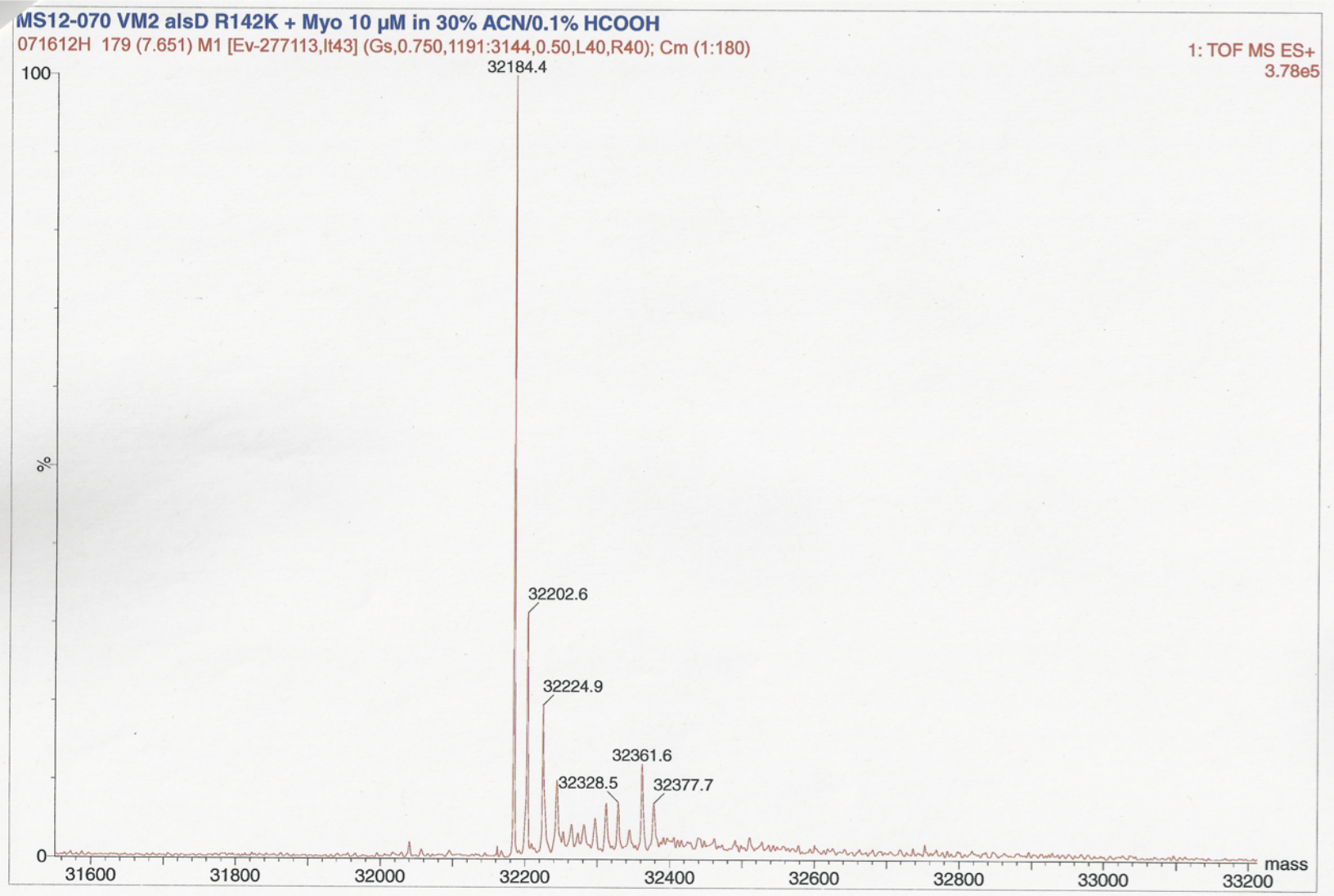


Figure D.9: Mass spectrum of recombinant R142A AlsD

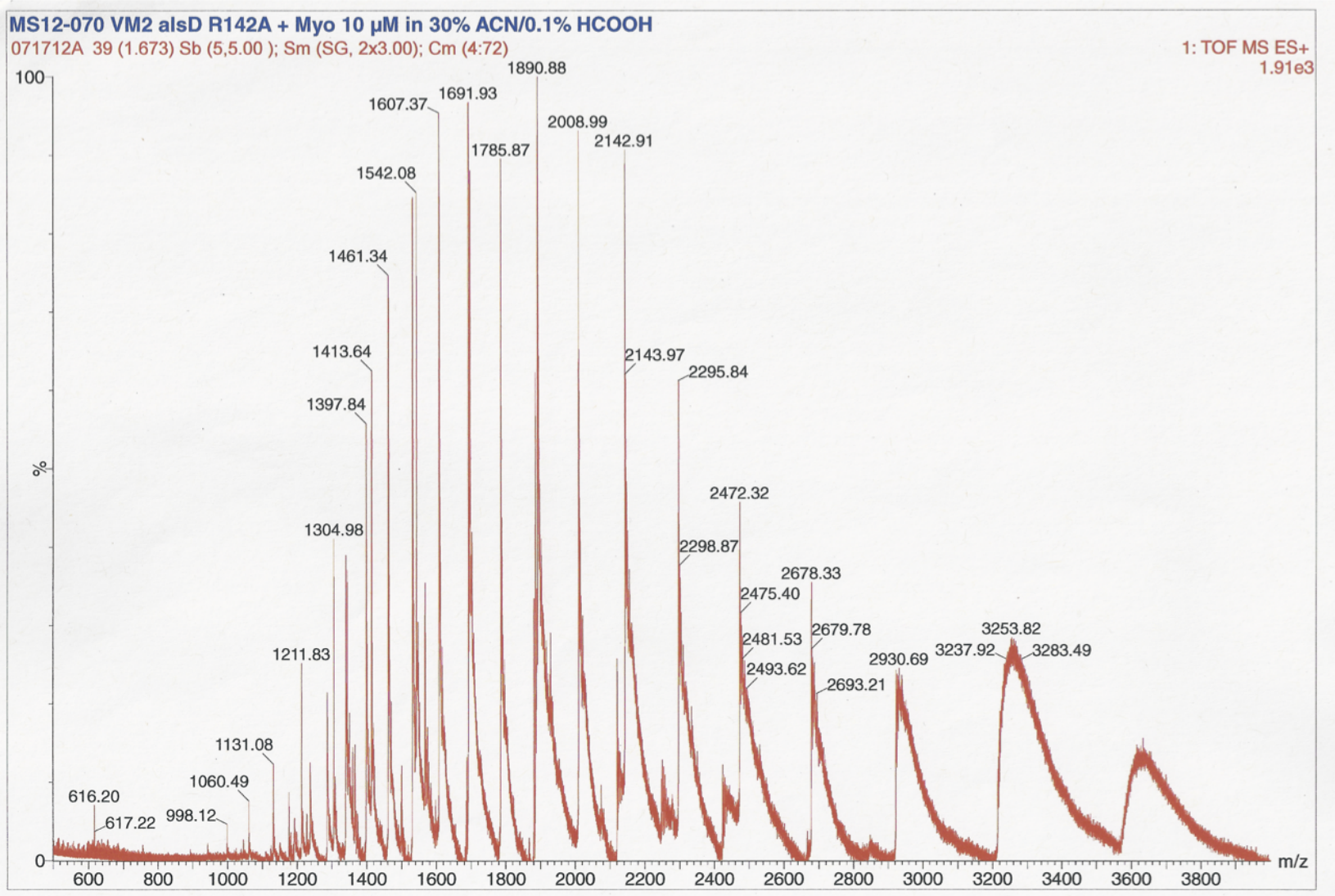


Figure D.10: Deconvoluted mass spectrum of recombinant R142A AlsD

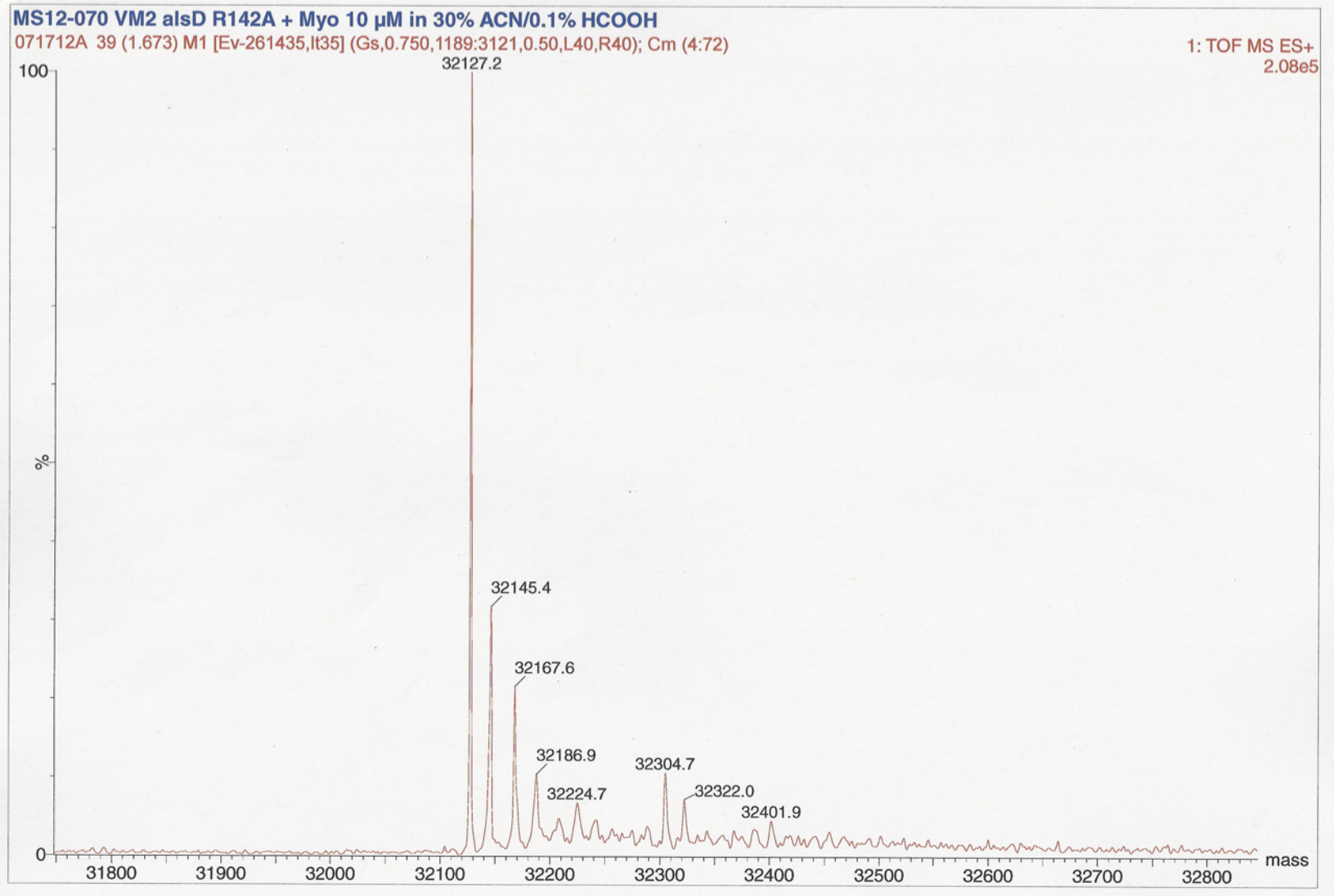


Figure D.11: Mass spectrum of recombinant E251Q AlsD

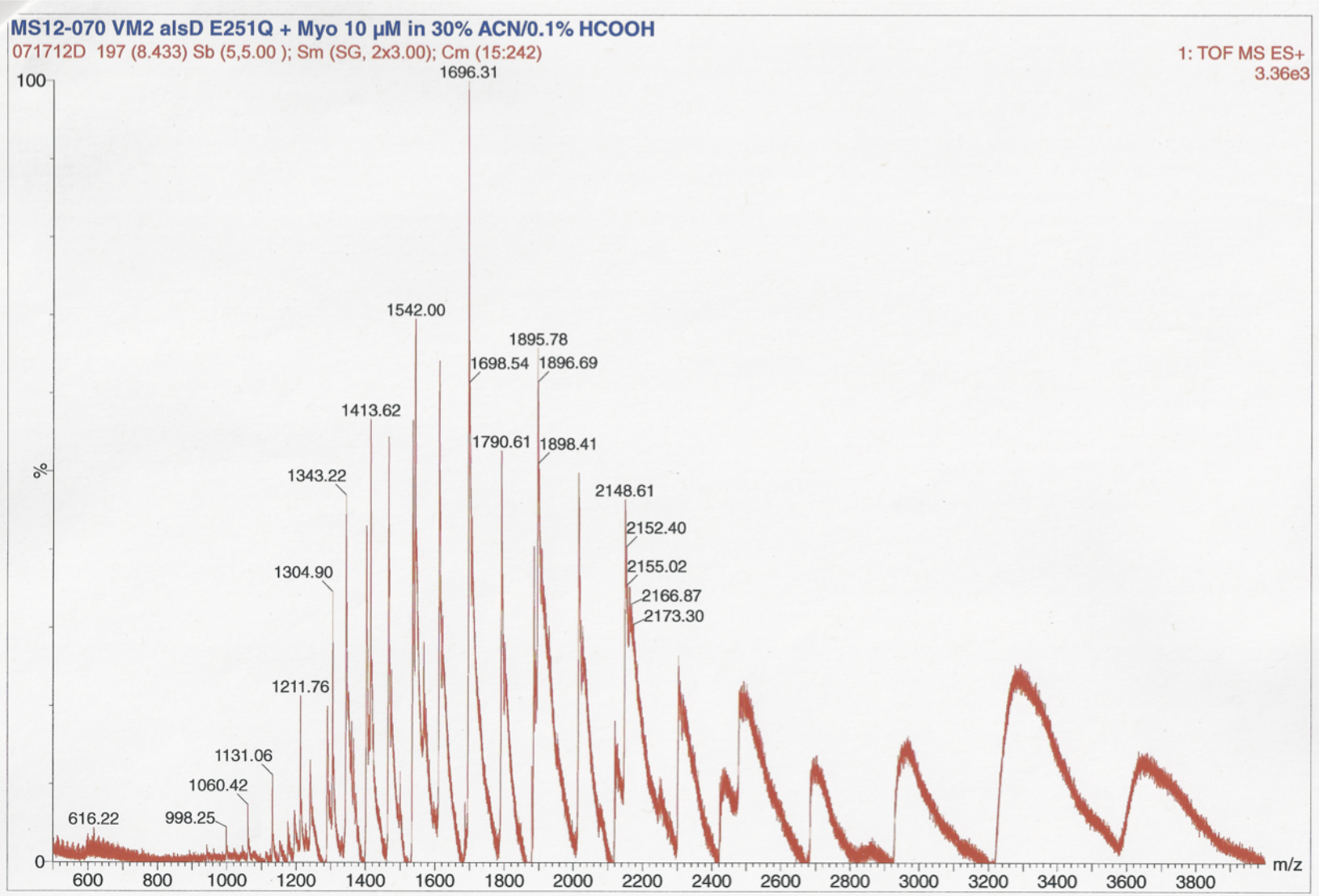


Figure D.12: Deconvoluted mass spectrum of recombinant E251Q AlsD

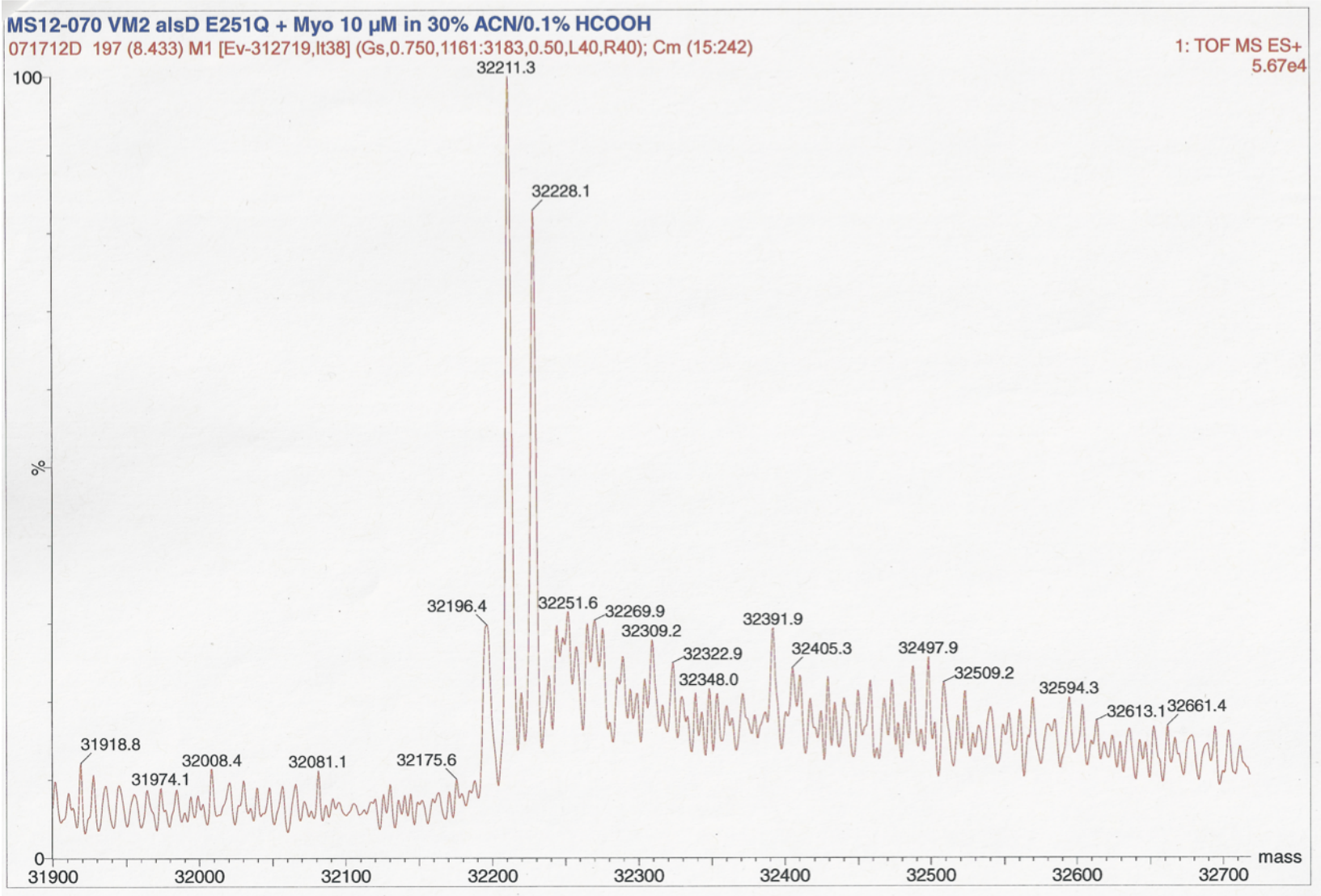


Figure D.13: Mass spectrum of recombinant E251A AlsD

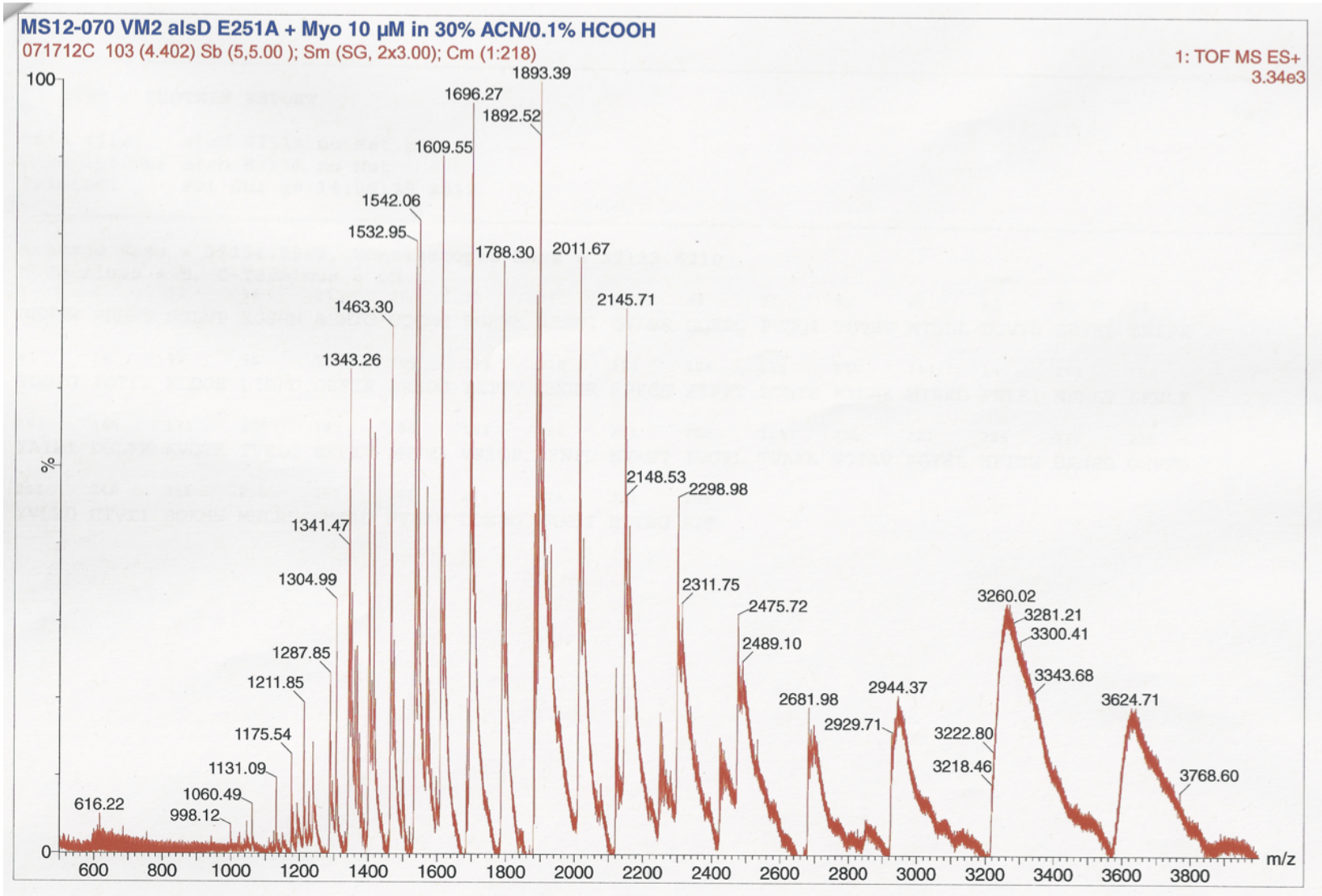


Figure D.14: Deconvoluted mass spectrum of recombinant E251A AlsD

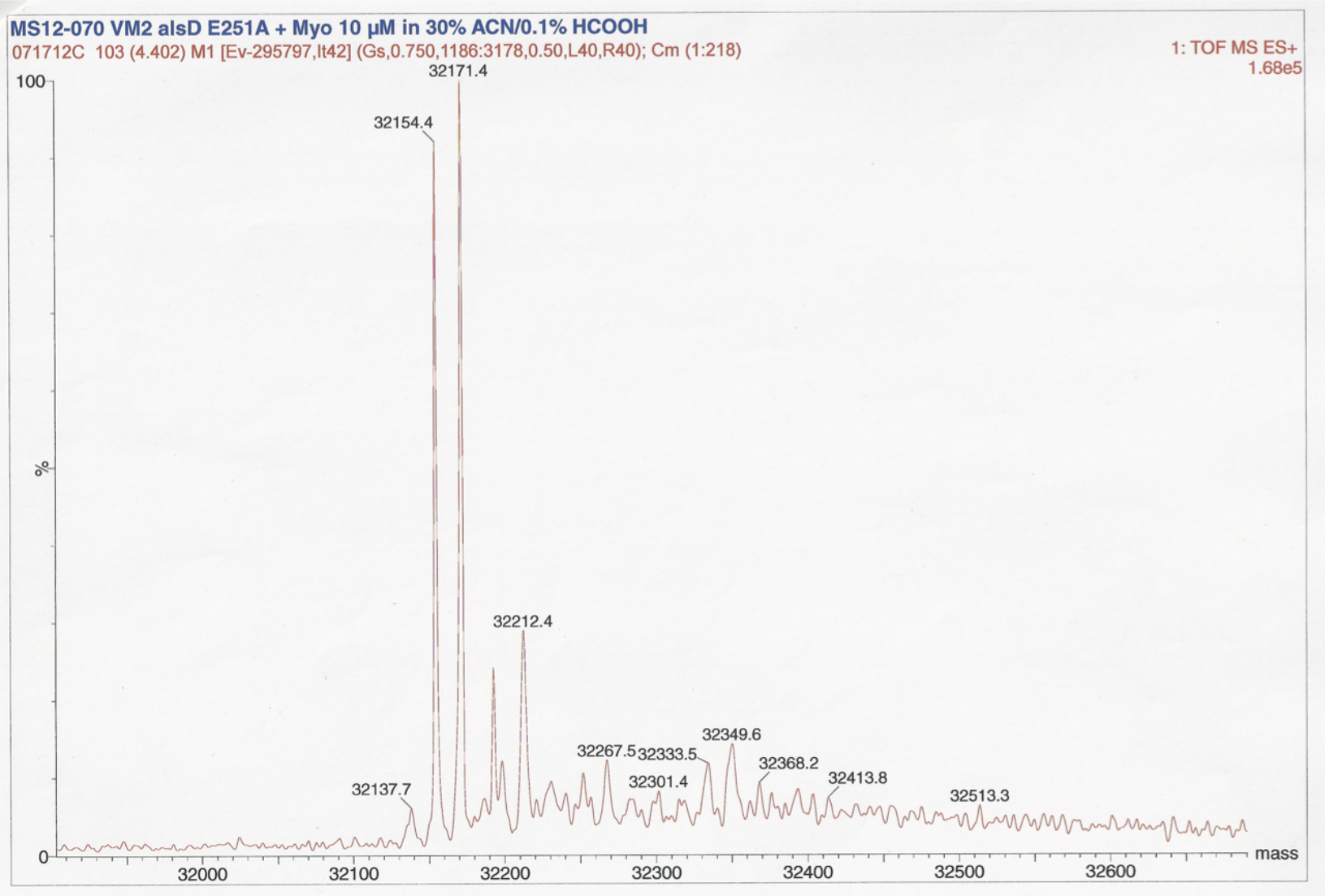


Figure D.15: Mass spectrum of recombinant T55S AlsD

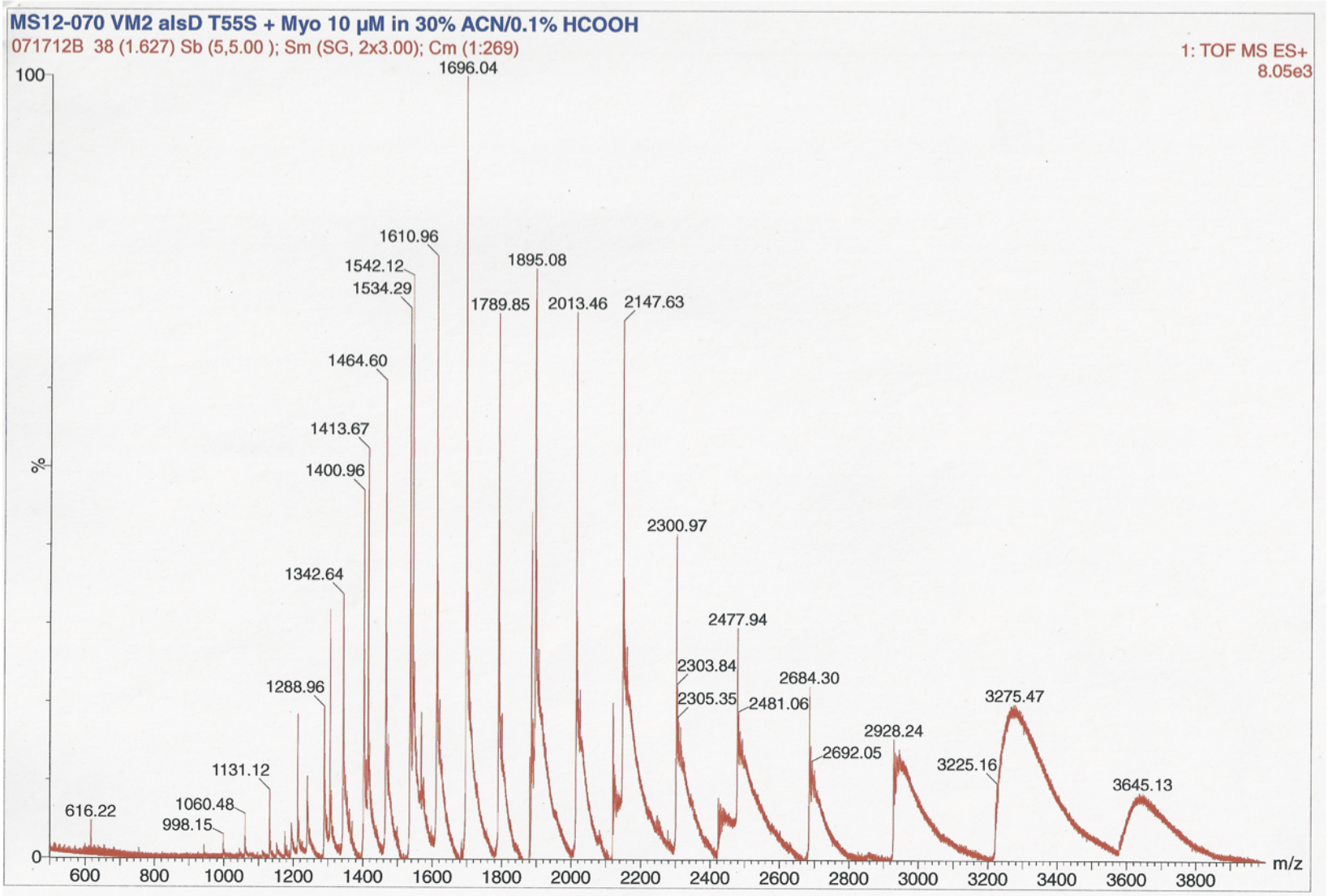


Figure D.16: Deconvoluted mass spectrum of recombinant T55S AlsD

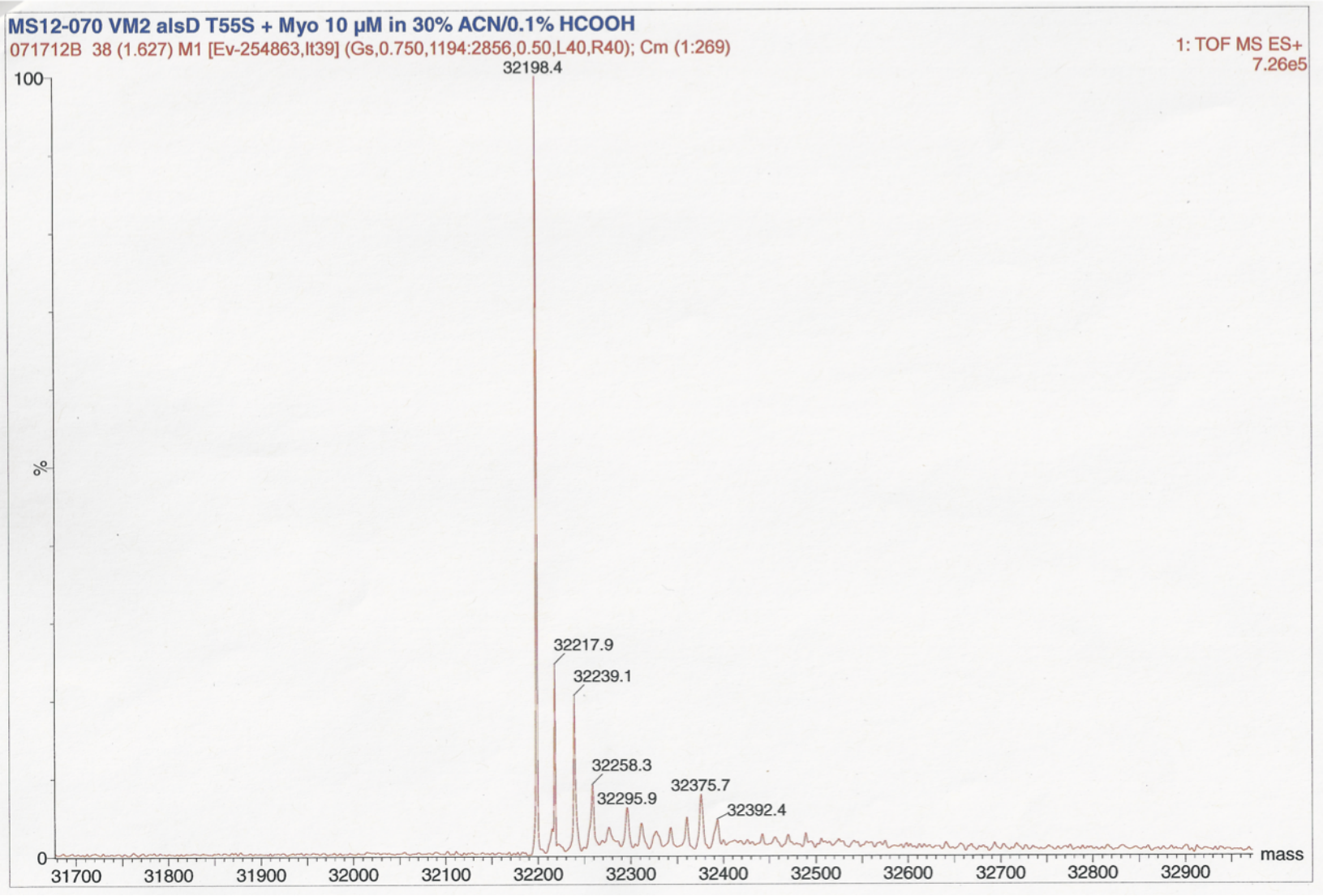


Figure D.17: Mass spectrum of recombinant T55A AlsD

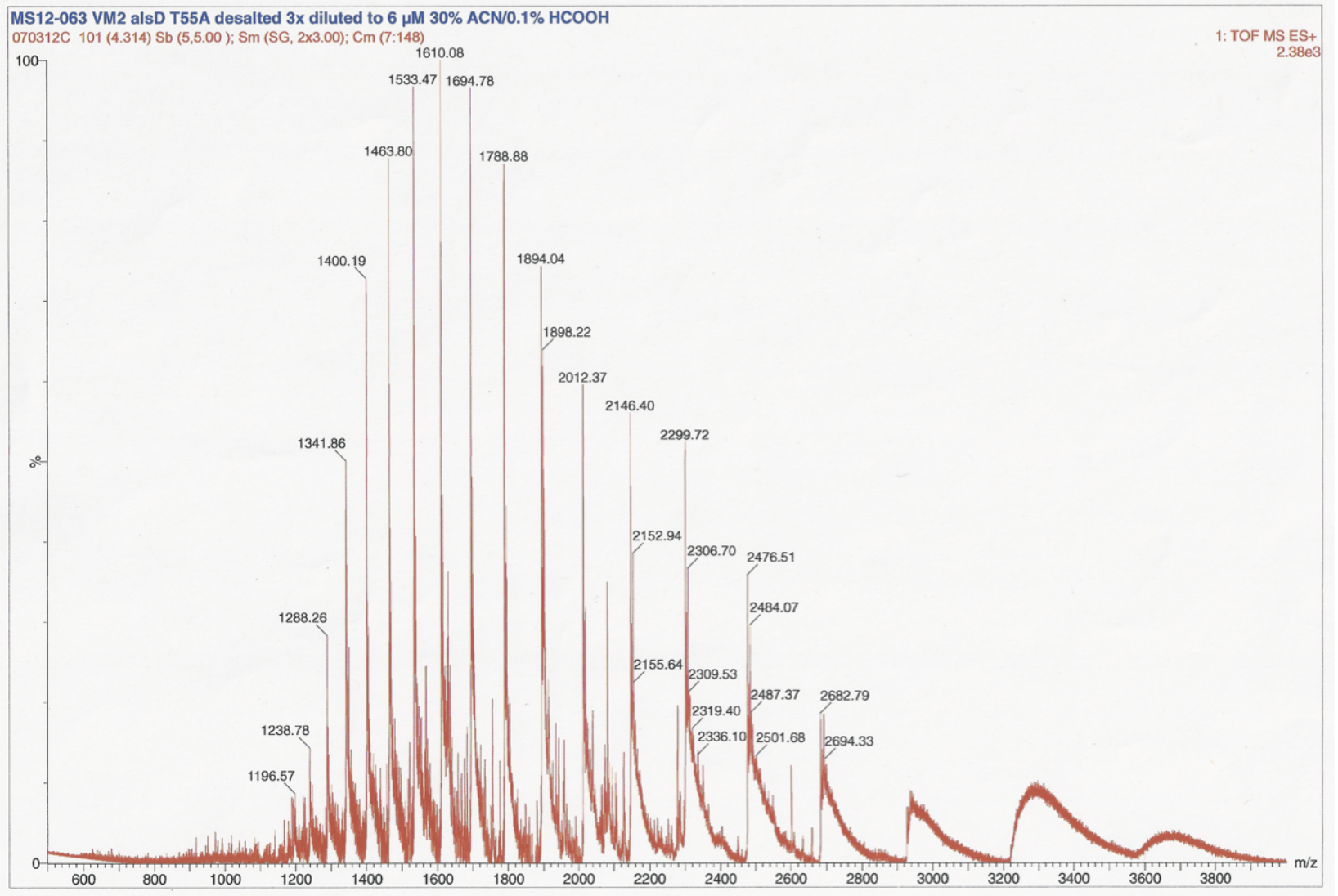
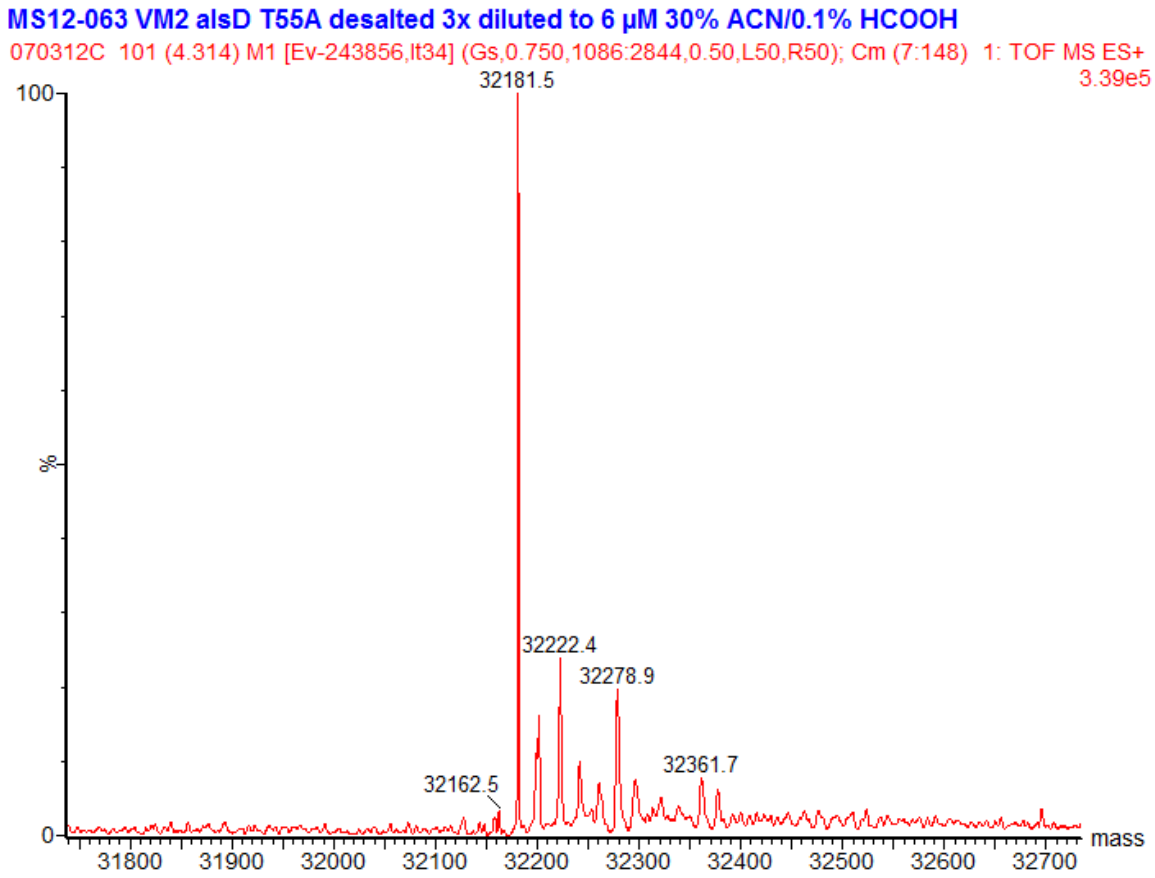


Figure D.18: Deconvoluted mass spectrum of recombinant T55A AlsD



Appendix E

CD spectra

This appendix contains the circular dichroism spectra obtained for *Bacillus brevis* ALDC and the recombinant *Bacillus subtilis* wild-type and mutant ALDC that were produced as described in Chapter 2 Methods and experimental. The figures show the experimentally measured spectra and the dichroweb reconstructed data.

Figure E.1: CD spectrum of *Bacillus brevis* ALDC

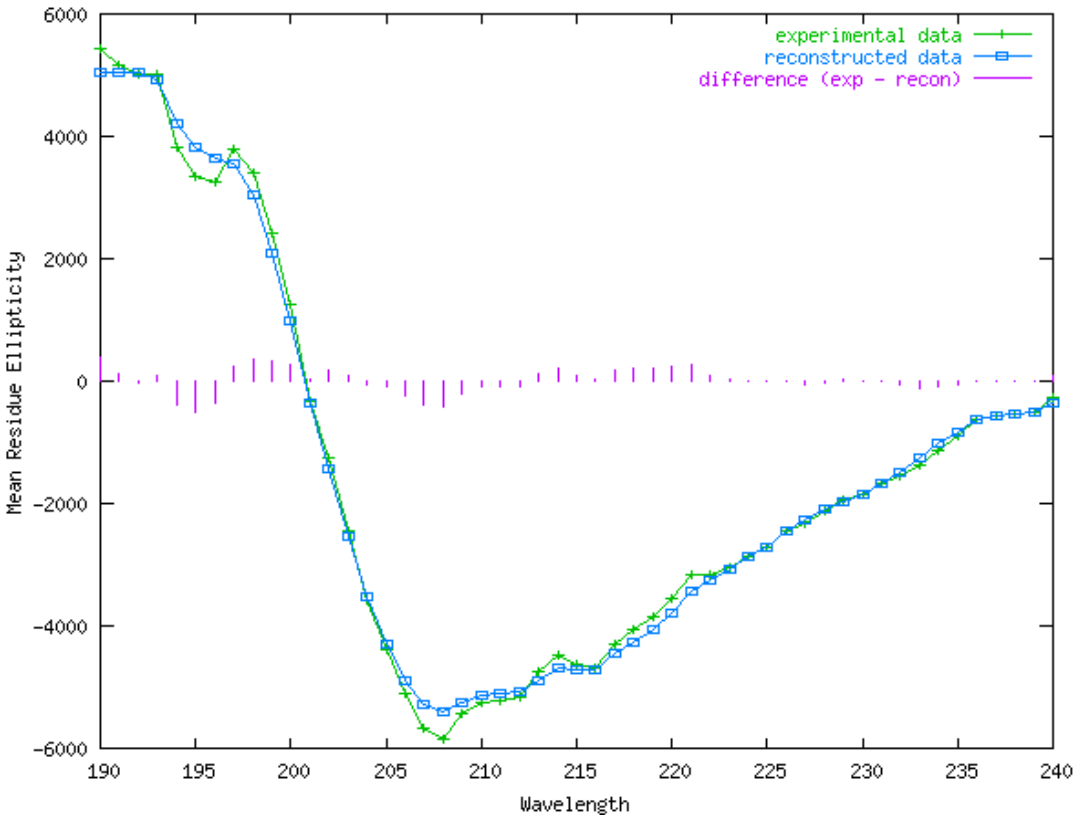


Figure E.2: CD spectrum of recombinant Bacillus subtilis wild-type AlsD

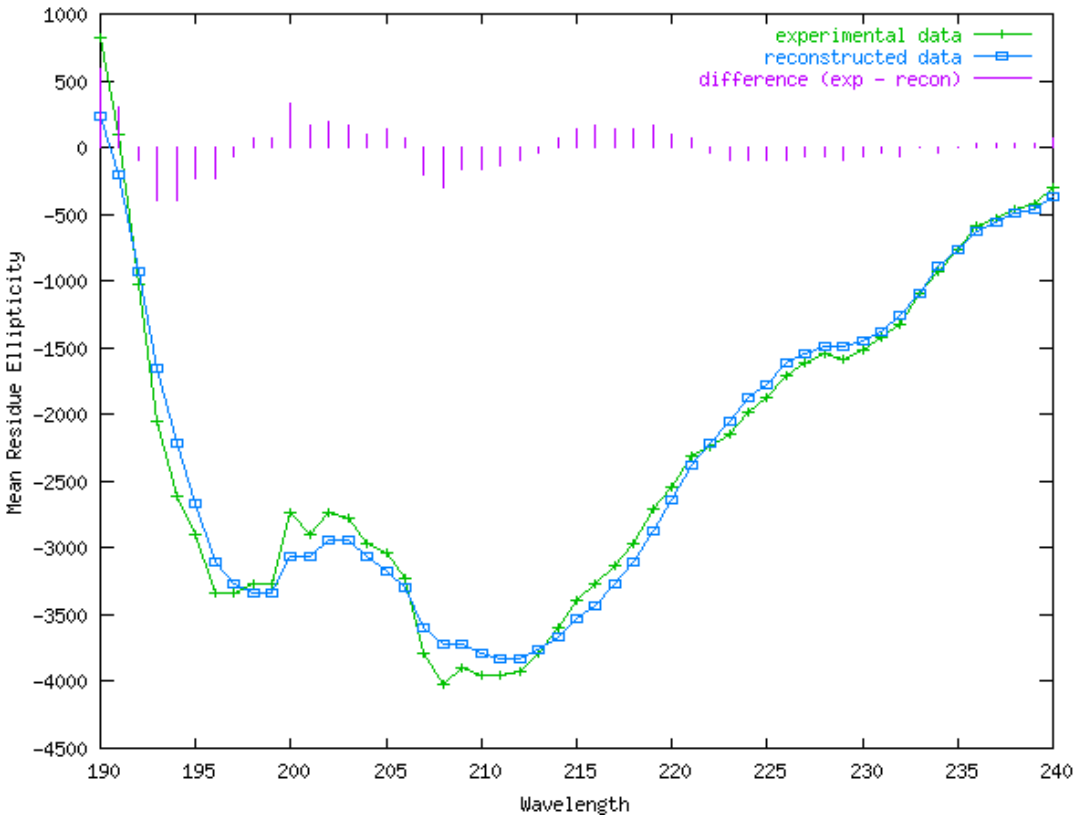


Figure E.3: CD spectrum of recombinant *Bacillus subtilis* E62Q AlsD

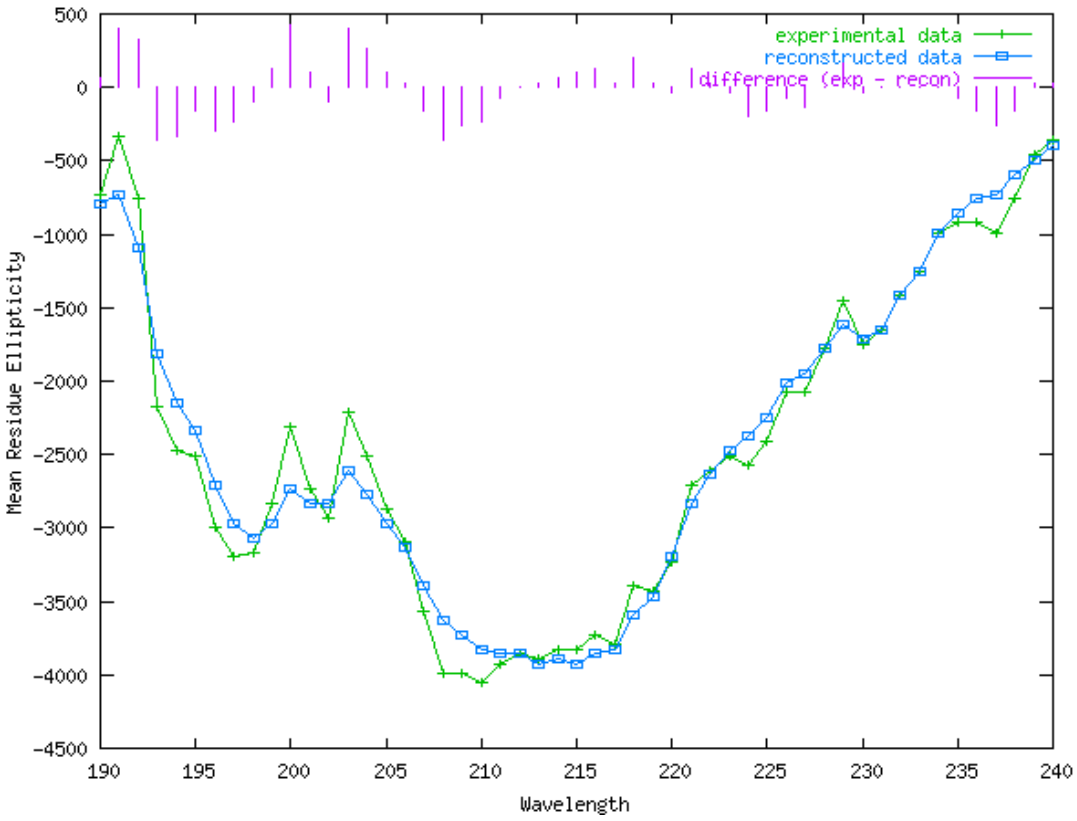


Figure E.4: CD spectrum of recombinant *Bacillus subtilis* E62A AlsD

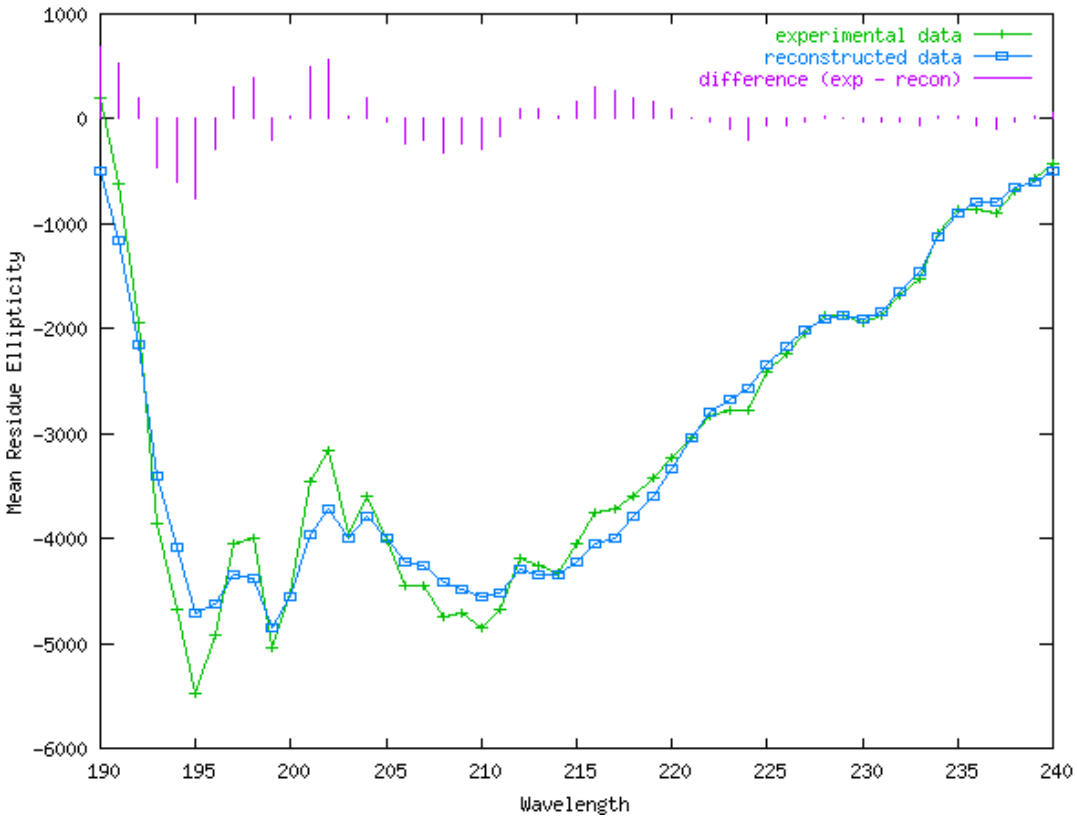


Figure E.5: CD spectrum of recombinant *Bacillus subtilis* R142K AlsD

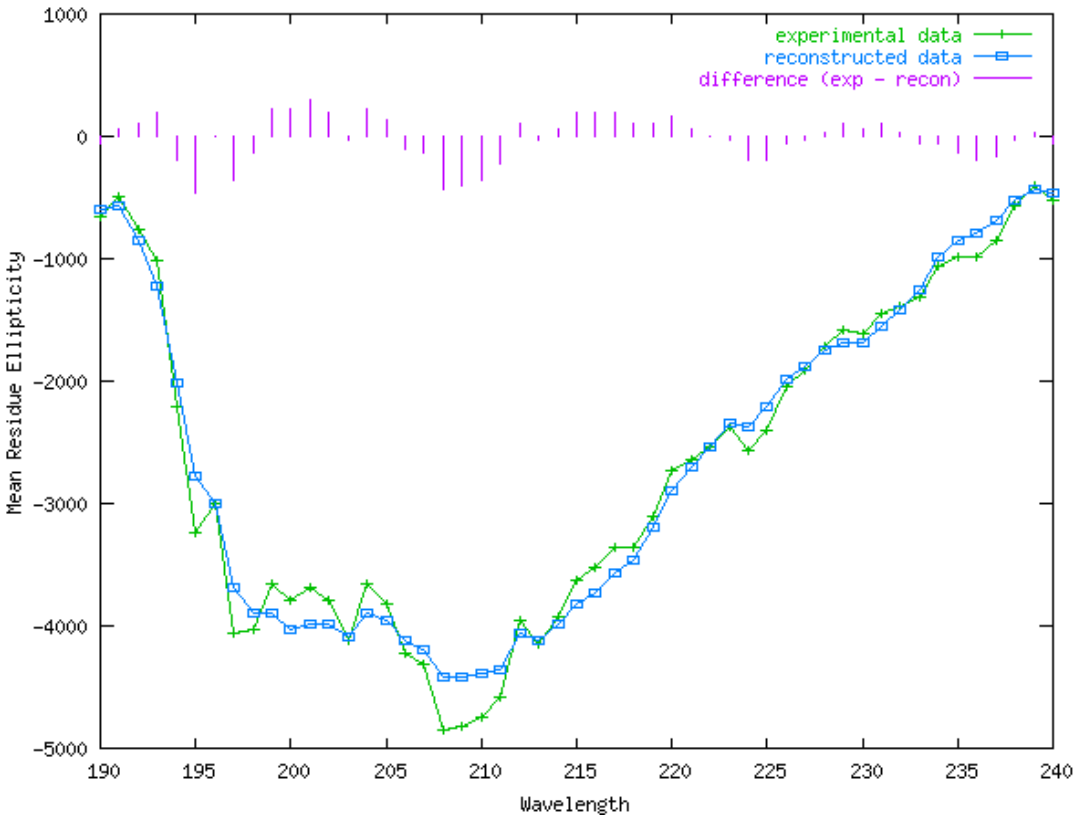


Figure E.6: CD spectrum of recombinant *Bacillus subtilis* R142A AlsD

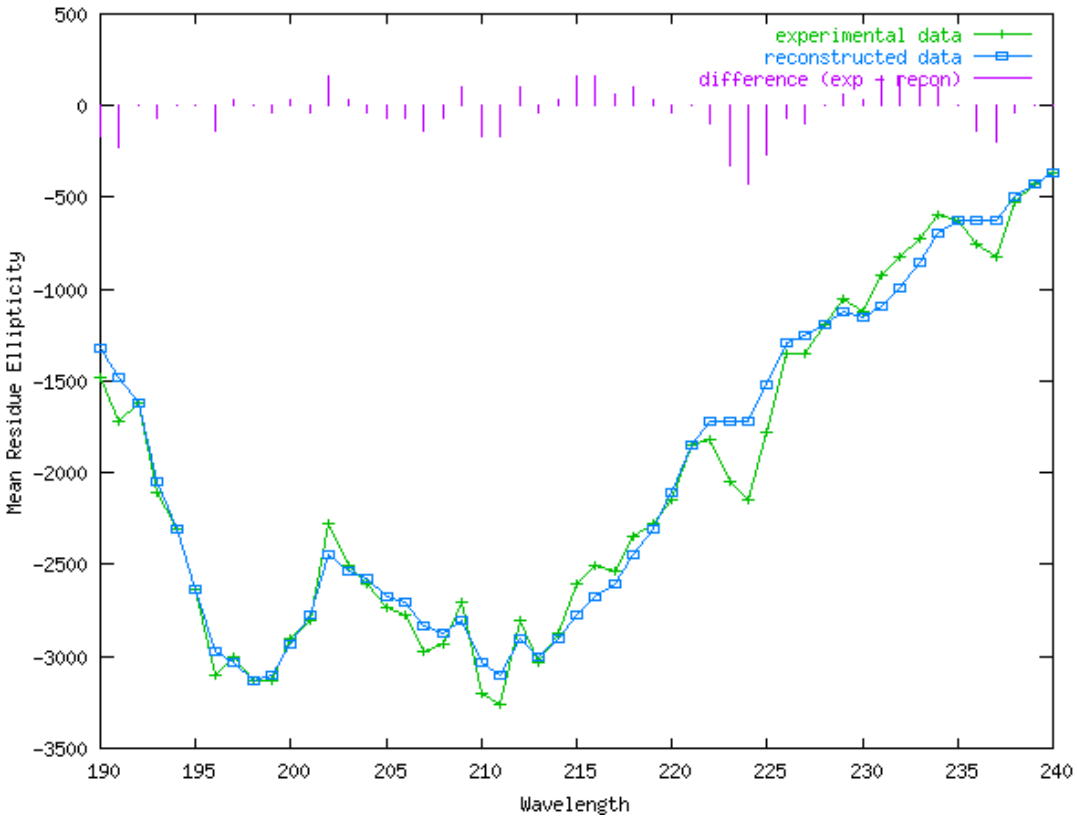


Figure E.7: CD spectrum of recombinant *Bacillus subtilis* E251Q AlsD

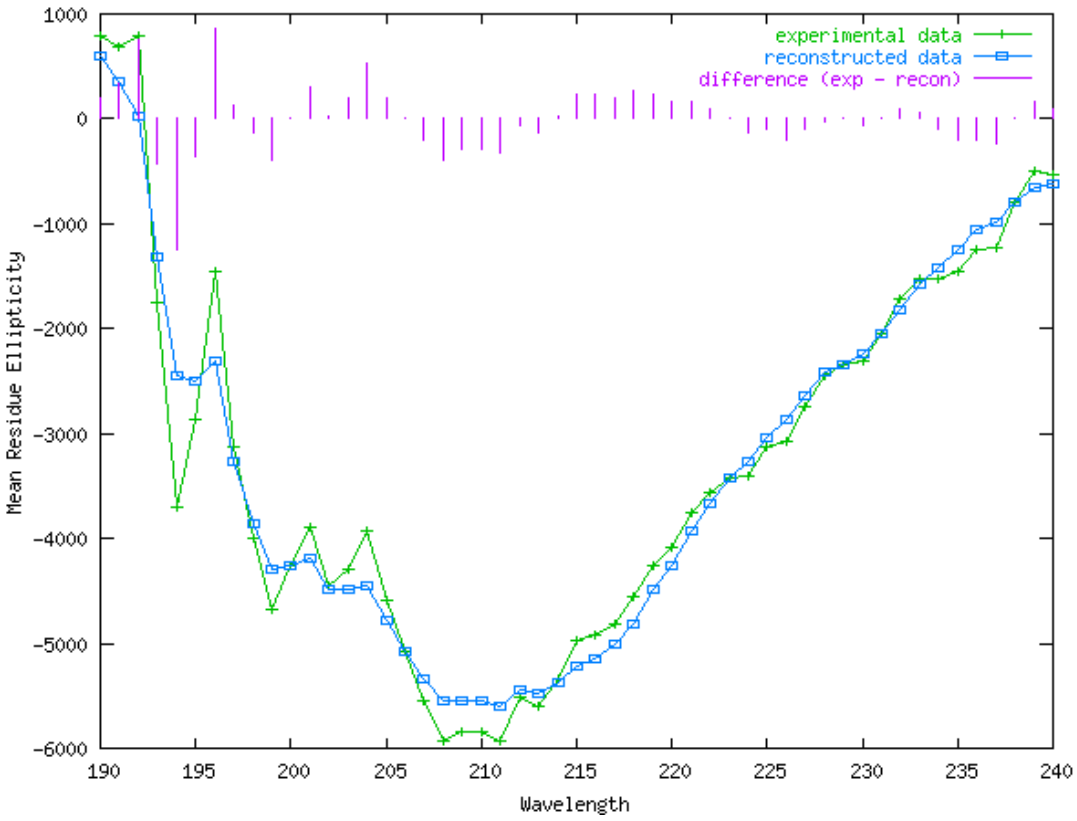


Figure E.8: CD spectrum of recombinant *Bacillus subtilis* E251A AlsD

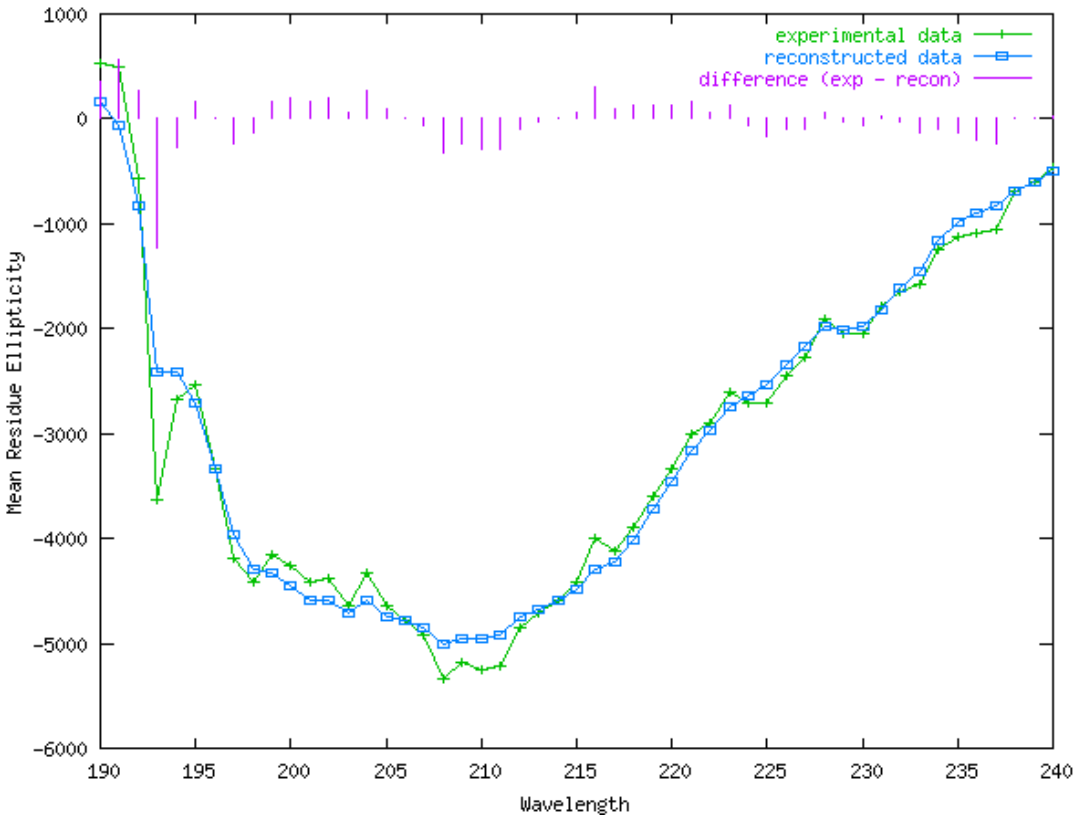


Figure E.9: CD spectrum of recombinant *Bacillus subtilis* T55S AlsD

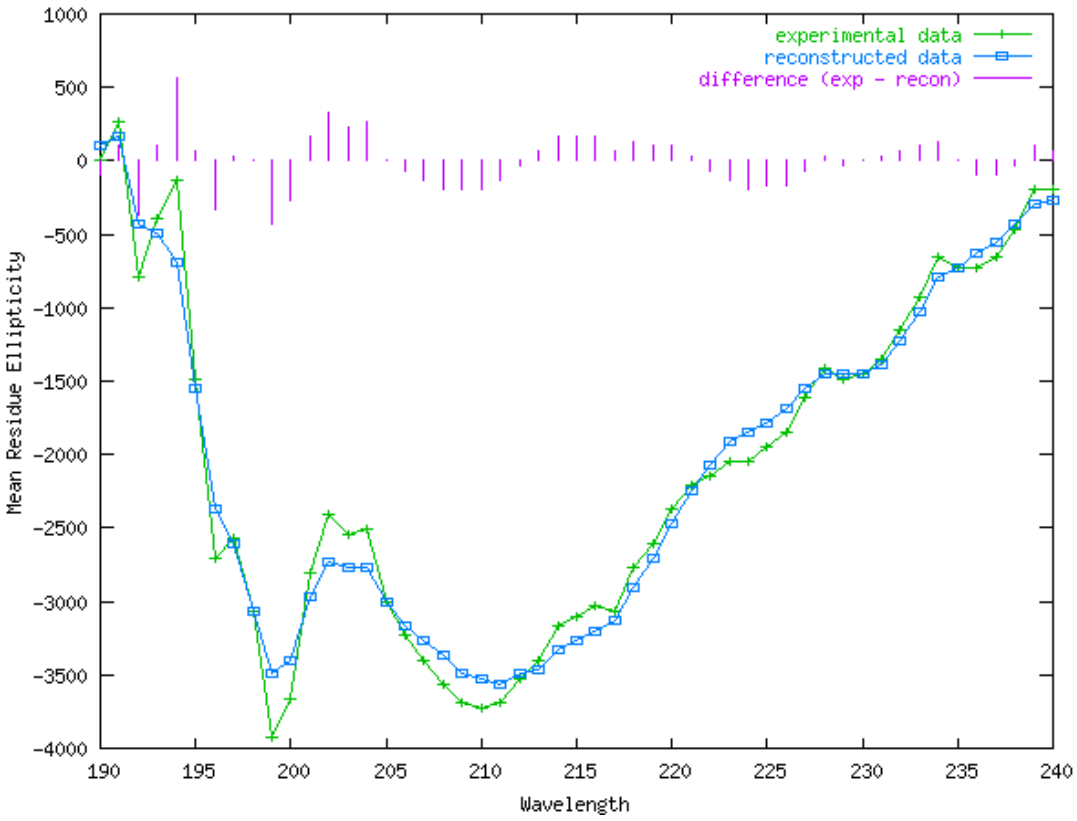
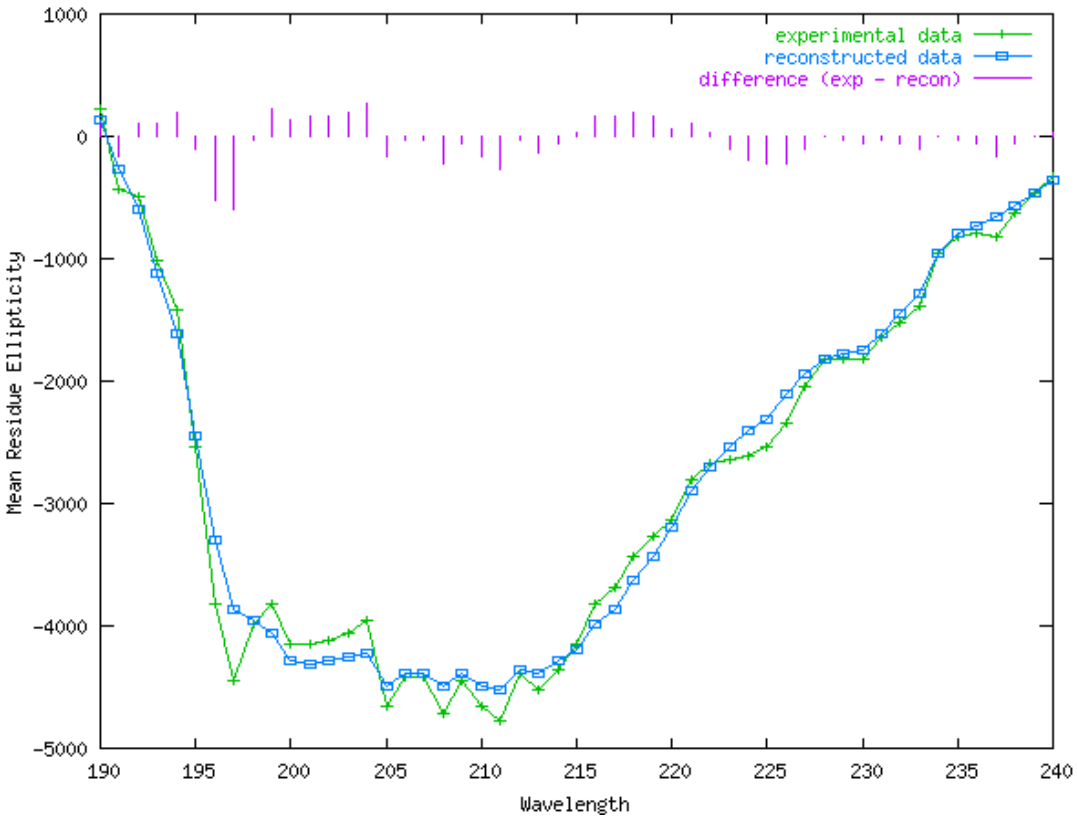


Figure E.10: CD spectrum of recombinant *Bacillus subtilis* T55A AlsD



Appendix F

CD calculations

F.1 Background

Circular dichroism is the difference in the absorbance of left circularly polarised light (A_L) and right circularly polarised light (A_R) as defined by Equation F.1.

$$\Delta A = A_L - A_R \tag{F.1}$$

The Jasco CD spectrometer used in this study outputs CD measurements in mdeg which are historic units of ellipticity (θ). Ellipticity can be converted into ΔA by the relationship described in Equation F.2.

$$\Delta A = \frac{\theta}{32982} \tag{F.2}$$

The Beer-Lambert law (Equation F.3) can be applied to CD and by rearranging the equation (Equation F.4) the molar CD ($\Delta\varepsilon$) can be calculated if the concentration and pathlength are

known.

$$\Delta A = \Delta\varepsilon \times c \times l \tag{F.3}$$

c = concentration (M), l = pathlength (cm)

$$\Delta\varepsilon = \frac{\Delta A}{c \times l} \tag{F.4}$$

The molar CD of (*S*)-acetolactate and (*R*)-acetoin were experimentally measured as described in Section 2.3.4.3 and used to convert from rates in mdeg/s into more meaningful rates for enzyme kinetic studies.

F.2 Conversion of units

Enzyme rates were measured, whilst linear, in units of mdeg s^{-1}. For kinetic analysis it was necessary to convert into more meaningful units of specific activity (μmol mg^{-1} min^{-1}). This was done in three stages, initially rates were converted into units of ΔA s^{-1} using the relationship described in F.2.

$$\text{rate } (\Delta\text{A s}^{-1}) = \frac{\text{rate (mdeg s}^{-1})}{32982} \tag{F.5}$$

Then rates were converted into units of concentration over time (mM s^{-1}) using the experimentally determined molar CD ($\Delta\varepsilon$), as shown in equation F.6.

$$\text{rate (mM s}^{-1}) = \frac{\text{rate } (\Delta\text{A s}^{-1})}{\frac{\Delta\varepsilon}{1000} \times l} \tag{F.6}$$

$\Delta\varepsilon$ = molar CD (M^{-1} cm^{-1}), l = pathlength (cm)

Equation F.7 shows how rates in units of mM s^{-1} were converted to specific activity in units of μmole of product per mg of enzyme per min.

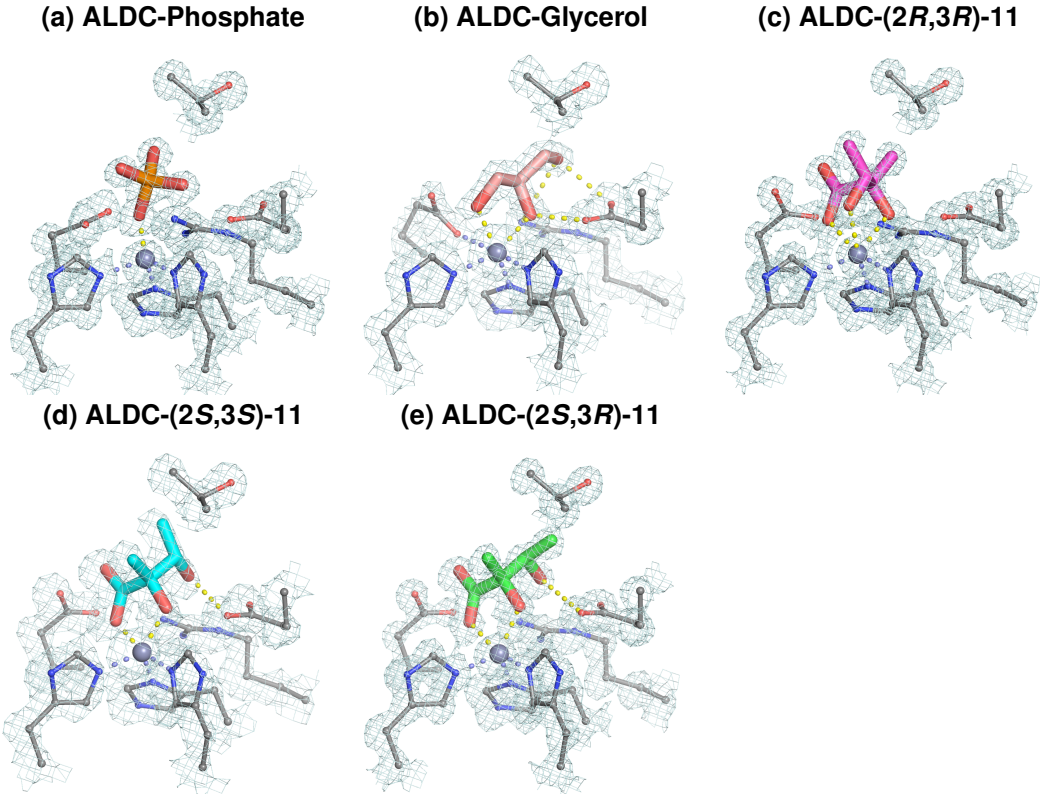
$$\text{Specific activity } (\mu\text{mol mg}^{-1}\ \text{min}^{-1}) = \frac{\text{rate (mM s}^{-1}) \times v_1\ \text{(ml)} \times 60\ \text{(s min}^{-1})}{c\ \text{(mg ml}^{-1}) \times v_2\ \text{(ml)}} \qquad \text{(F.7)}$$

v_1 = total assay volume (ml), c = protein concentration (mg mg^{-1}), v_2 = volume of protein added (ml)

Appendix G

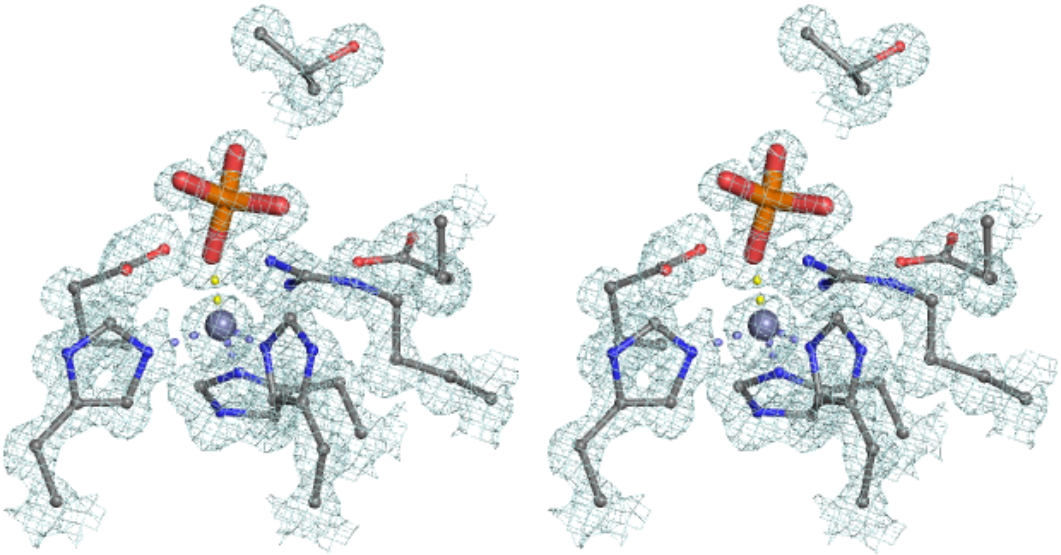
B. brevis ALDC active site structures

Figure G.1: Structures of the ALDC active site



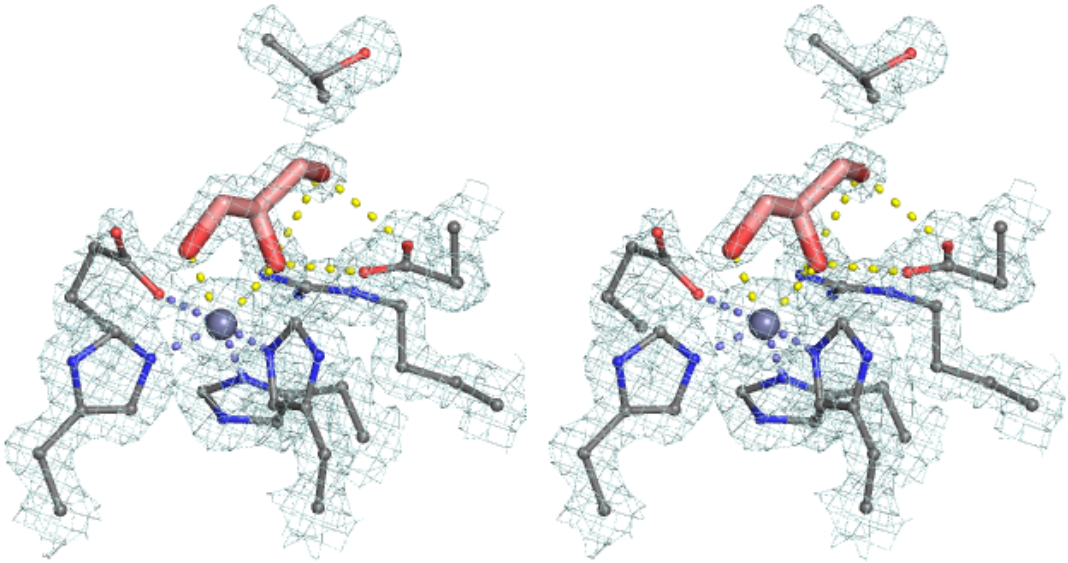
*Structures of ALDC in complex with (a) phosphate, (b) glycerol, (c) (2R,3R)-**11**, (d) (2S,3S)-**11** and (e) (2S,3R)-**11**. Electron density (2mFo-ΔFc) maps are shown as a blue mesh, they are contoured at 1.5 σ level and contours more than 1.4 Å from any displayed atoms have been removed for clarity.*

Figure G.2: Stereo view image of the ALDC-Phosphate structure



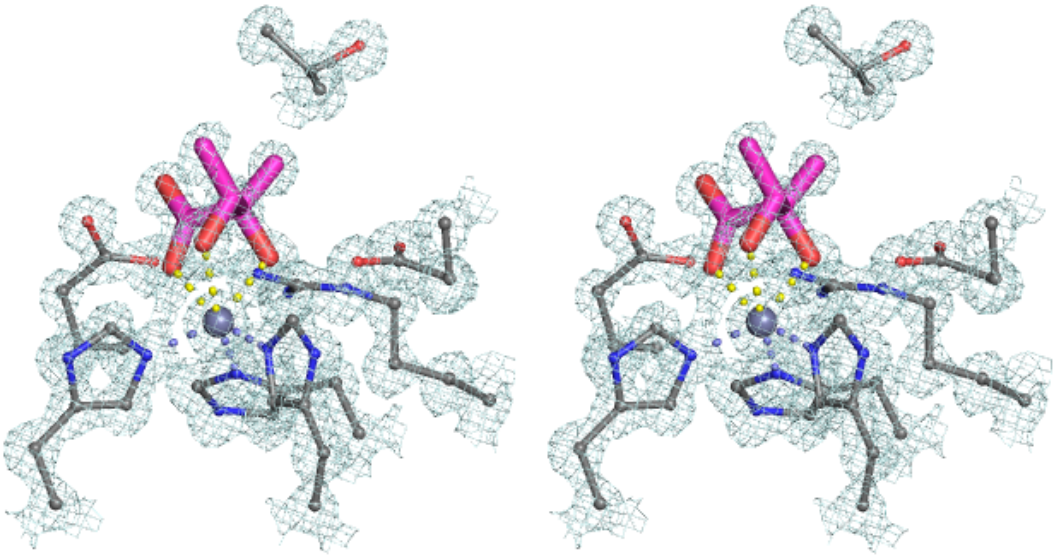
Stereo view figure of the ALDC-Phosphate active site structure generated using the wall-eyed stereo view setting in Pymol. The electron density map was generated as described in Figure G.1.

Figure G.3: Stereo view image of the ALDC-Glycerol structure



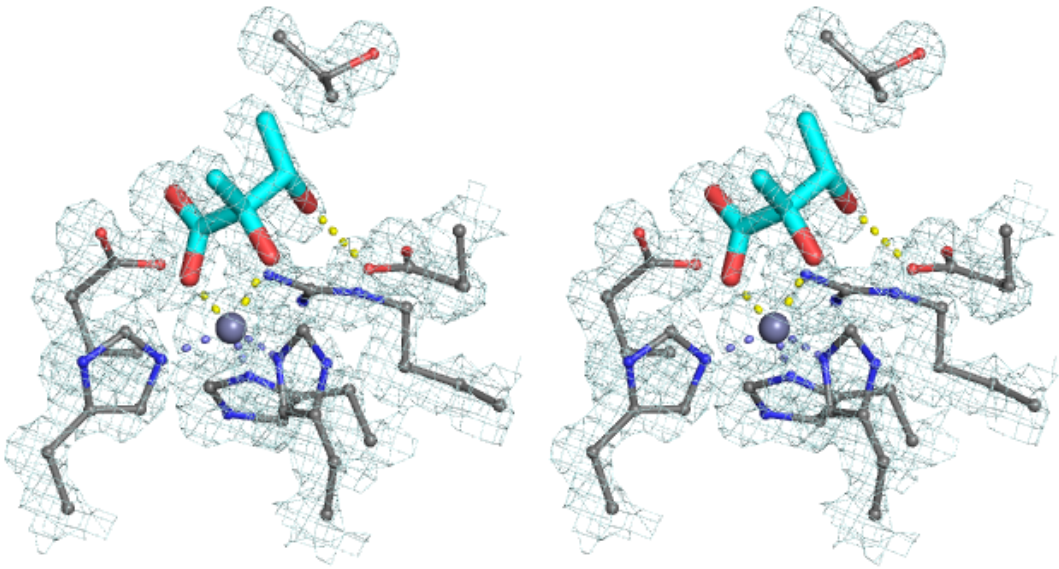
Stereo view figure of the ALDC-Glycerol active site structure generated using the wall-eyed stereo view setting in Pymol. The electron density map was generated as described in Figure G.1.

Figure G.4: Stereo view image of the ALDC-(2*R*,3*R*)-11 structure



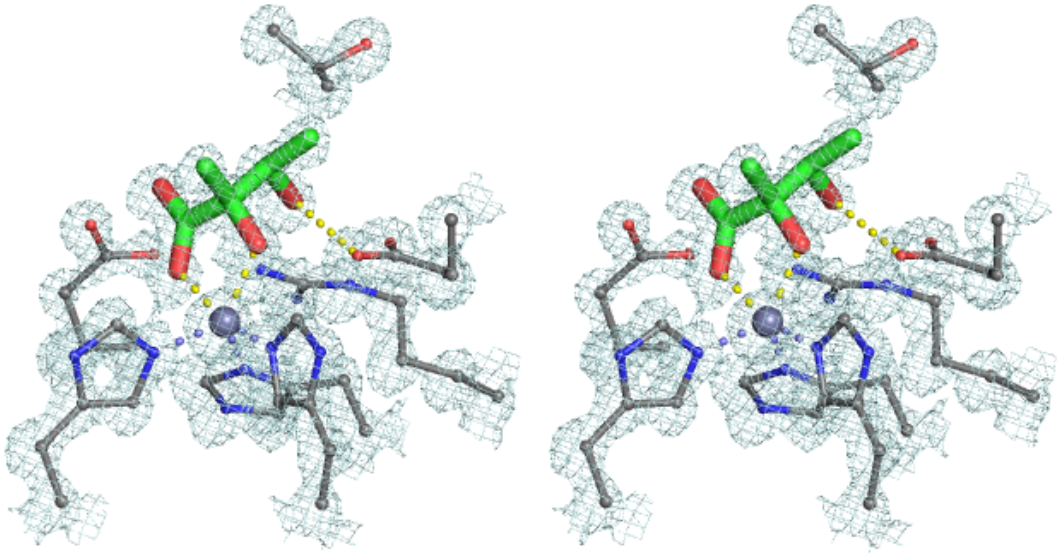
*Stereo view figure of the ALDC-(2R,3R)-**11** active site structure generated using the wall-eyed stereo view setting in Pymol. The electron density map was generated as described in Figure G.1.*

Figure G.5: Stereo view image of the ALDC-(2*S*,3*S*)-11 structure



*Stereo view figure of the ALDC-(2S,3S)-**11** active site structure generated using the wall-eyed stereo view setting in Pymol. The electron density map was generated as described in Figure G.1.*

Figure G.6: Stereo view image of the ALDC-(2*S*,3*R*)-11 structure



*Stereo view figure of the ALDC-(2S,3R)-**11** active site structure generated using the wall-eyed stereo view setting in Pymol. The electron density map was generated as described in Figure G.1.*